

PCTWORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁷ : C12N 15/12, C07K 14/47, C12N 15/62, C07K 16/18		A2	(11) International Publication Number: WO 00/15796 (43) International Publication Date: 23 March 2000 (23.03.00)
(21) International Application Number: PCT/US99/21090 (22) International Filing Date: 15 September 1999 (15.09.99) (30) Priority Data: PCT/US98/19330 16 September 1998 (16.09.98) US		(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).	
(71) Applicant (<i>for all designated States except US</i>): GENENTECH, INC. [US/US]; 1 DNA Way, South San Francisco, CA 94080-4990 (US). (72) Inventors; and (75) Inventors/Applicants (<i>for US only</i>): CHEN, Jian [CN/US]; 121 York Drive, Princeton, NJ 08540 (US). GODDARD, Audrey [CA/US]; 110 Congo Street, San Francisco, CA 94131 (US). GURNEY, Austin, L. [US/US]; 1 Debbie Lane, Belmont, CA 94002 (US). HILLAN, Kenneth [GB/US]; 64 Seward Street, San Francisco, CA 94114 (US). PENNICA, Diane [US/US]; 2417 Hale Drive, Burlingame, CA 94010 (US). WOOD, William, I. [US/US]; 35 Southdown Court, Hillsborough, CA 94010 (US). YUAN, Jean [CN/US]; 176 West 37th Avenue, San Mateo, CA 94003 (US). (74) Agents: KRESNAK, Mark, T. et al.; Genentech Inc., 1 DNA Way, South San Francisco, CA 94080-4990 (US).		Published <i>Without international search report and to be republished upon receipt of that report.</i>	
(54) Title: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC ACIDS ENCODING THE SAME			
(57) Abstract			
<p>The present invention is directed to novel polypeptides and to nucleic acid molecules encoding those polypeptides. Also provided herein are vectors and host cells comprising those nucleic acid sequences, chimeric polypeptide molecules comprising the polypeptides of the present invention fused to heterologous polypeptide sequences, antibodies which bind to the polypeptides of the present invention and to methods for producing the polypeptides of the present invention.</p>			

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC ACIDS ENCODING THE
SAME

FIELD OF THE INVENTION

The present invention relates generally to the identification and isolation of novel DNA and to the
5 recombinant production of novel polypeptides.

BACKGROUND OF THE INVENTION

Extracellular proteins play important roles in, among other things, the formation, differentiation and maintenance of multicellular organisms. The fate of many individual cells, e.g., proliferation, migration, 10 differentiation, or interaction with other cells, is typically governed by information received from other cells and/or the immediate environment. This information is often transmitted by secreted polypeptides (for instance, mitogenic factors, survival factors, cytotoxic factors, differentiation factors, neuropeptides, and hormones) which are, in turn, received and interpreted by diverse cell receptors or membrane-bound proteins. These secreted polypeptides or signaling molecules normally pass through the cellular secretory pathway to reach their site of 15 action in the extracellular environment.

Secreted proteins have various industrial applications, including as pharmaceuticals, diagnostics, biosensors and bioreactors. Most protein drugs available at present, such as thrombolytic agents, interferons, interleukins, erythropoietins, colony stimulating factors, and various other cytokines, are secretory proteins. Their receptors, which are membrane proteins, also have potential as therapeutic or diagnostic agents. Efforts 20 are being undertaken by both industry and academia to identify new, native secreted proteins. Many efforts are focused on the screening of mammalian recombinant DNA libraries to identify the coding sequences for novel secreted proteins. Examples of screening methods and techniques are described in the literature [see, for example, Klein et al., *Proc. Natl. Acad. Sci.* 93:7108-7113 (1996); U.S. Patent No. 5,536,637].

Membrane-bound proteins and receptors can play important roles in, among other things, the formation, 25 differentiation and maintenance of multicellular organisms. The fate of many individual cells, e.g., proliferation, migration, differentiation, or interaction with other cells, is typically governed by information received from other cells and/or the immediate environment. This information is often transmitted by secreted polypeptides (for instance, mitogenic factors, survival factors, cytotoxic factors, differentiation factors, neuropeptides, and hormones) which are, in turn, received and interpreted by diverse cell receptors or membrane-bound proteins. 30 Such membrane-bound proteins and cell receptors include, but are not limited to, cytokine receptors, receptor kinases, receptor phosphatases, receptors involved in cell-cell interactions, and cellular adhesin molecules like selectins and integrins. For instance, transduction of signals that regulate cell growth and differentiation is regulated in part by phosphorylation of various cellular proteins. Protein tyrosine kinases, enzymes that catalyze that process, can also act as growth factor receptors. Examples include fibroblast growth factor receptor and 35 nerve growth factor receptor.

Membrane-bound proteins and receptor molecules have various industrial applications, including as pharmaceutical and diagnostic agents. Receptor immunoadhesins, for instance, can be employed as therapeutic agents to block receptor-ligand interactions. The membrane-bound proteins can also be employed for screening of potential peptide or small molecule inhibitors of the relevant receptor/ligand interaction.

Efforts are being undertaken by both industry and academia to identify new, native receptor or 5 membrane-bound proteins. Many efforts are focused on the screening of mammalian recombinant DNA libraries to identify the coding sequences for novel receptor or membrane-bound proteins.

1. PRO211 and PRO217

Epidermal growth factor (EGF) is a conventional mitogenic factor that stimulates the proliferation of 10 various types of cells including epithelial cells and fibroblasts. EGF binds to and activates the EGF receptor (EGFR), which initiates intracellular signaling and subsequent effects. The EGFR is expressed in neurons of the cerebral cortex, cerebellum, and hippocampus in addition to other regions of the central nervous system (CNS). In addition, EGF is also expressed in various regions of the CNS. Therefore, EGF acts not only on mitotic cells, but also on postmitotic neurons. In fact, many studies have indicated that EGF has neurotrophic 15 or neuromodulatory effects on various types of neurons in the CNS. For example, EGF acts directly on cultured cerebral cortical and cerebellar neurons, enhancing neurite outgrowth and survival. On the other hand, EGF also acts on other cell types, including septal cholinergic and mesencephalic dopaminergic neurons, indirectly through glial cells. Evidence of the effects of EGF on neurons in the CNS is accumulating, but the mechanisms of action remain essentially unknown. EGF-induced signaling in mitotic cells is better understood than in 20 postmitotic neurons. Studies of cloned pheochromocytoma PC12 cells and cultured cerebral cortical neurons have suggested that the EGF-induced neurotrophic actions are mediated by sustained activation of the EGFR and mitogen-activated protein kinase (MAPK) in response to EGF. The sustained intracellular signaling correlates with the decreased rate of EGFR down-regulation, which might determine the response of neuronal cells to EGF. It is likely that EGF is a multi-potent growth factor that acts upon various types of cells including mitotic cells 25 and postmitotic neurons.

EGF is produced by the salivary and Brunner's glands of the gastrointestinal system, kidney, pancreas, thyroid gland, pituitary gland, and the nervous system, and is found in body fluids such as saliva, blood, cerebrospinal fluid (CSF), urine, amniotic fluid, prostatic fluid, pancreatic juice, and breast milk, *Plata-Salaman, Peptides 12: 653-663 (1991)*.

30 EGF is mediated by its membrane specific receptor, which contains an intrinsic tyrosine kinase. Stoscheck *et al.*, *J. Cell Biochem. 31: 135-152 (1986)*. EGF is believed to function by binding to the extracellular portion of its receptor which induces a transmembrane signal that activates the intrinsic tyrosine kinase.

Purification and sequence analysis of the EGF-like domain has revealed the presence of six conserved 35 cysteine residues which cross-bind to create three peptide loops, Savage *et al.*, *J. Biol. Chem. 248: 7669-7672 (1979)*. It is now generally known that several other peptides can react with the EGF receptor which share the same generalized motif $X_nCX_2CX_{4/5}CX_{10}CXCX_5GX_2CX_n$, where X represents any non-cysteine amino acid, and

n is a variable repeat number. Non isolated peptides having this motif include TGF- α , amphiregulin, schwannoma-derived growth factor (SDGF), heparin-binding EGF-like growth factors and certain virally encoded peptides (e.g., Vaccinia virus, Reisner, *Nature* 313: 801-803 (1985), Shope fibroma virus, Chang et al., *Mol Cell Biol.* 7: 535-540 (1987), Molluscum contagiosum, Porter and Archard, *J. Gen. Virol.* 68: 673-682 (1987), and Myxoma virus, Upton et al., *J. Virol.* 61: 1271-1275 (1987), Prigent and Lemoine, *Prog. Growth*

5 *Factor Res.* 4: 1-24 (1992).

EGF-like domains are not confined to growth factors but have been observed in a variety of cell-surface and extracellular proteins which have interesting properties in cell adhesion, protein-protein interaction and development, Laurence and Gusterson, *Tumor Biol.* 11: 229-261 (1990). These proteins include blood coagulation factors (factors VI, IX, X, XII, protein C, protein S, protein Z, tissue plasminogen activator, 10 urokinase), extracellular matrix components (laminin, cytactin, entactin), cell surface receptors (LDL receptor, thrombomodulin receptor) and immunity-related proteins (complement C1r, uromodulin).

Even more interesting, the general structure pattern of EGF-like precursors is preserved through lower organisms as well as in mammalian cells. A number of genes with developmental significance have been identified in invertebrates with EGF-like repeats. For example, the *notch* gene of *Drosophila* encodes 15 tandemly arranged 40 amino acid repeats which show homology to EGF, Wharton et al., *Cell* 43: 557-581 (1985). Hydropathy plots indicate a putative membrane spanning domain, with the EGF-related sequences being located on the extracellular side of the membrane. Other homeotic genes with EGF-like repeats include Delta, 95F and 5ZD which were identified using probes based on Notch, and the nematode gene *Lin-12* which encodes a putative receptor for a developmental signal transmitted between two specified cells.

20 Specifically, EGF has been shown to have potential in the preservation and maintenance of gastrointestinal mucosa and the repair of acute and chronic mucosal lesions, Konturek et al., *Eur. J. Gastroenterol Hepatol.* 7 (10), 933-37 (1995), including the treatment of necrotizing enterocolitis, Zollinger-Ellison syndrome, gastrointestinal ulceration gastrointestinal ulcerations and congenital microvillus atrophy, Guglietta and Sullivan, *Eur. J. Gastroenterol Hepatol.* 7(10), 945-50 (1995). Additionally, EGF has been 25 implicated in hair follicle differentiation; du Cros, *J. Invest. Dermatol.* 101 (1 Suppl.), 106S-113S (1993), Hillier, *Clin. Endocrinol.* 93(4), 427-28 (1990); kidney function, Hamm et al., *Semin. Nephrol.* 13 (1): 109-15 (1993), Harris, *Am. J. Kidney Dis.* 17(6): 627-30 (1991); tear fluid, van Setten et al., *Int. Ophthalmol.* 15(6): 359-62 (1991); vitamin K mediated blood coagulation, Stenflo et al., *Blood* 78(7): 1637-51 (1991). EGF is also implicated various skin disease characterized by abnormal keratinocyte differentiation, e.g., psoriasis, epithelial 30 cancers such as squamous cell carcinomas of the lung, epidermoid carcinoma of the vulva and gliomas. King et al., *Am. J. Med. Sci.* 296: 154-158 (1988).

Of great interest is mounting evidence that genetic alterations in growth factors signaling pathways are closely linked to developmental abnormalities and to chronic diseases including cancer. Aaronson, *Science* 254: 1146-1153 (1991). For example, c-erb-2 (also known as HER-2), a proto-oncogene with close structural 35 similarity to EGF receptor protein, is overexpressed in human breast cancer. King et al., *Science* 229: 974-976 (1985); Gullick, *Hormones and their actions*, Cooke et al., eds, Amsterdam, Elsevier, pp 349-360 (1986).

We herein describe the identification and characterization of novel polypeptides having homology to EGF, wherein those polypeptides are herein designated PRO211 and PRO217.

2. PRO230

Nephritis is a condition characterized by inflammation of the kidney affecting the structure and normal function of the kidney. This condition can be chronic or acute and is generally caused by infection, degenerative process or vascular disease. In all cases, early detection is desirable so that the patient with nephritis can begin treatment of the condition.

An approach to detecting nephritis is to determine the antigens associated with nephritis and antibodies thereto. In rabbit, a tubulointerstitial nephritis antigen (TIN-ag) has been reported in Nelson, T. R., et al., J. Biol. Chem., 270(27):16265-70 (July 1995) (GENBANK/U24270). This study reports that the rabbit TIN-ag is a basement membrane glycoprotein having a predicted amino acid sequence which has a carboxyl-terminal region exhibiting 30% homology with human procathepsin B, a member of the cystein proteinase family of proteins. It is also reported that the rabbit TIN-ag has a domain in the amino-terminal region containing an epidermal growth factor-like motif that shares homology with laminin A and S chains, alpha 1 chain of type I collagen, von Willebrand's factor and mucin, indicating structural and functional similarities. Studies have also been conducted in mice. However, it is desirable to identify tubulointerstitial nephritis antigens in humans to aid in the development of early detection methods and treatment of nephritis.

Proteins which have homology to tubulointerstitial nephritis antigens are of particular interest to the medical and industrial communities. Often, proteins having homology to each other have similar function. It is also of interest when proteins having homology do not have similar functions, indicating that certain structural motifs identify information other than function, such as locality of function. We herein describe the identification and characterization of a novel polypeptide, designated herein as PRO230, which has homology to tubulointerstitial nephritis antigens.

25 3. PRO232

Stem cells are undifferentiated cells capable of (a) proliferation, (b) self maintenance, (c) the production of a large number of differentiated functional progeny, (d) regeneration of tissue after injury and/or (e) a flexibility in the use of these options. Stem cells often express cell surface antigens which are capable of serving as cell specific markers that can be exploited to identify stem cells, thereby providing a means for identifying 30 and isolating specific stem cell populations.

Having possession of different stem cell populations will allow for a number of important applications. For example, possessing a specific stem cell population will allow for the identification of growth factors and other proteins which are involved in their proliferation and differentiation. In addition, there may be as yet undiscovered proteins which are associated with (1) the early steps of dedication of the stem cell to a particular 35 lineage, (2) prevention of such dedication, and (3) negative control of stem cell proliferation, all of which may be identified if one has possession of the stem cell population. Moreover, stem cells are important and ideal targets for gene therapy where the inserted genes promote the health of the individual into whom the stem cells

are transplanted. Finally, stem cells may play important roles in transplantation of organs or tissues, for example liver regeneration and skin grafting.

Given the importance of stem cells in various different applications, efforts are currently being undertaken by both industry and academia to identify new, native stem cell antigen proteins so as to provide specific cell surface markers for identifying stem cell populations as well as for providing insight into the 5 functional roles played by stem cell antigens in cell proliferation and differentiation. We herein describe the identification and characterization of novel polypeptides having homology to a stem cell antigen, wherein those polypeptides are herein designated as PRO232 polypeptides..

4. PRO187

10 Growth factors are molecular signals or mediators that enhance cell growth or proliferation, alone or in concert, by binding to specific cell surface receptors. However, there are other cellular reactions than only growth upon expression to growth factors. As a result, growth factors are better characterized as multifunctional and potent cellular regulators. Their biological effects include proliferation, chemotaxis and stimulation of extracellular matrix production. Growth factors can have both stimulatory and inhibitory effects. For example, 15 transforming growth factor (TGF- β) is highly pleiotropic and can stimulate proliferation in some cells, especially connective tissue, while being a potent inhibitor of proliferation in others, such as lymphocytes and epithelial cells.

20 The physiological effect of growth stimulation or inhibition by growth factors depends upon the state of development and differentiation of the target tissue. The mechanism of local cellular regulation by classical endocrine molecules involves comprehends autocrine (same cell), juxtacrine (neighbor cell), and paracrine (adjacent cells) pathways. Peptide growth factors are elements of a complex biological language, providing the basis for intercellular communication. They permit cells to convey information between each other, mediate interaction between cells and change gene expression. The effect of these multifunctional and pluripotent factors is dependent on the presence or absence of other peptides.

25 FGF-8 is a member of the fibroblast growth factors (FGFs) which are a family of heparin-binding, potent mitogens for both normal diploid fibroblasts and established cell lines, Gospodarowicz *et al.* (1984), *Proc. Natl. Acad. Sci. USA* 81:6963. The FGF family comprises acidic FGF (FGF-1), basic FGF (FGF-2), INT-2 (FGF-3), K-FGF/HST (FGF-4), FGF-5, FGF-6, KGF (FGF-7), AIGF (FGF-8) among others. All FGFs have two conserved cysteine residues and share 30-50% sequence homology at the amino acid level. These factors 30 are mitogenic for a wide variety of normal diploid mesoderm-derived and neural crest-derived cells, including granulosa cells, adrenal cortical cells, chondrocytes, myoblasts, corneal and vascular endothelial cells (bovine or human), vascular smooth muscle cells, lens, retina and prostatic epithelial cells, oligodendrocytes, astrocytes, chondrocytes, myoblasts and osteoblasts.

35 Fibroblast growth factors can also stimulate a large number of cell types in a non-mitogenic manner. These activities include promotion of cell migration into wound area (chemotaxis), initiation of new blood vessel formulation (angiogenesis), modulation of nerve regeneration and survival (neurotrophism), modulation of endocrine functions, and stimulation or suppression of specific cellular protein expression, extracellular matrix

production and cell survival. Baird & Bohlen, *Handbook of Exp. Pharmacol.* 95(1): 369-418, Springer, (1990). These properties provide a basis for using fibroblast growth factors in therapeutic approaches to accelerate wound healing, nerve repair, collateral blood vessel formation, and the like. For example, fibroblast growth factors have been suggested to minimize myocardium damage in heart disease and surgery (U.S.P. 4,378,347).

5 FGF-8, also known as androgen-induced growth factor (AIGF), is a 215 amino acid protein which shares 30-40% sequence homology with the other members of the FGF family. FGF-8 has been proposed to be under androgenic regulation and induction in the mouse mammary carcinoma cell line SC3. Tanaka *et al.*, *Proc. Natl. Acad. Sci. USA* 89: 8928-8932 (1992); Sato *et al.*, *J. Steroid Biochem. Molec. Biol.* 47: 91-98 (1993). As a result, FGF-8 may have a local role in the prostate, which is known to be an androgen-responsive organ. FGF-8 can also be oncogenic, as it displays transforming activity when transfected into NIH-3T3 10 fibroblasts. Kouhara *et al.*, *Oncogene* 9: 455-462 (1994). While FGF-8 has been detected in heart, brain, lung, kidney, testis, prostate and ovary, expression was also detected in the absence of exogenous androgens. Schmitt *et al.*, *J. Steroid Biochem. Mol. Biol.* 57 (3-4): 173-78 (1996).

15 FGF-8 shares the property with several other FGFs of being expressed at a variety of stages of murine embryogenesis, which supports the theory that the various FGFs have multiple and perhaps coordinated roles in differentiation and embryogenesis. Moreover, FGF-8 has also been identified as a protooncogene that cooperates with Wnt-1 in the process of mammary tumorigenesis (Shackleford *et al.*, *Proc. Natl. Acad. Sci. USA* 90, 740-744 (1993); Heikinheimo *et al.*, *Mech. Dev.* 48: 129-138 (1994)).

20 In contrast to the other FGFs, FGF-8 exists as three protein isoforms, as a result of alternative splicing of the primary transcript. Tanaka *et al.*, *supra*. Normal adult expression of FGF-8 is weak and confined to gonadal tissue, however northern blot analysis has indicated that FGF-8 mRNA is present from day 10 through 25 day 12 or murine gestation, which suggests that FGF-8 is important to normal development. Heikinheimo *et al.*, *Mech Dev.* 48(2): 129-38 (1994). Further *in situ* hybridization assays between day 8 and 16 of gestation indicated initial expression in the surface ectoderm of the first bronchial arches, the frontonasal process, the forebrain and the midbrain-hindbrain junction. At days 10-12, FGF-8 was expressed in the surface ectoderm of the forelimb and hindlimb buds, the nasal pits and nasopharynx, the infundibulum and in the telencephalon, diencephalon and metencephalon. Expression continues in the developing hindlimbs through day 13 of gestation, but is undetectable thereafter. The results suggest that FGF-8 has a unique temporal and spatial pattern in embryogenesis and suggests a role for this growth factor in multiple regions of ectodermal differentiation in the post-gastrulation embryo.

30 We herein describe the identification of novel polypeptides having homology to FGF-8, wherein those polypeptides are herein designated PRO187 polypeptides.

5. PRO265

35 Protein-protein interactions include receptor and antigen complexes and signaling mechanisms. As more is known about the structural and functional mechanisms underlying protein-protein interactions, protein-protein interactions can be more easily manipulated to regulate the particular result of the protein-protein interaction. Thus, the underlying mechanisms of protein-protein interactions are of interest to the scientific and medical

community.

All proteins containing leucine-rich repeats are thought to be involved in protein-protein interactions. Leucine-rich repeats are short sequence motifs present in a number of proteins with diverse functions and cellular locations. The crystal structure of ribonuclease inhibitor protein has revealed that leucine-rich repeats correspond to beta-alpha structural units. These units are arranged so that they form a parallel beta-sheet with one surface exposed to solvent, so that the protein acquires an unusual, nonglobular shape. These two features have been indicated as responsible for the protein-binding functions of proteins containing leucine-rich repeats. See, Kobe and Deisenhofer, Trends Biochem. Sci., 19(10):415-421 (Oct. 1994).

A study has been reported on leucine-rich proteoglycans which serve as tissue organizers, orienting and ordering collagen fibrils during ontogeny and are involved in pathological processes such as wound healing, tissue repair, and tumor stroma formation. Iozzo, R. V., Crit. Rev. Biochem. Mol. Biol., 32(2):141-174 (1997). Other studies implicating leucine rich proteins in wound healing and tissue repair are De La Salle, C., et al., Vouv. Rev. Fr. Hematol. (Germany), 37(4):215-222 (1995), reporting mutations in the leucine rich motif in a complex associated with the bleeding disorder Bernard-Soulier syndrome and Chlemetson, K. J., Thromb. Haemost. (Germany), 74(1):111-116 (July 1995), reporting that platelets have leucine rich repeats. Another protein of particular interest which has been reported to have leucine-rich repeats is the SLIT protein which has been reported to be useful in treating neuro-degenerative diseases such as Alzheimer's disease, nerve damage such as in Parkinson's disease, and for diagnosis of cancer, see, Artavanistsakonas, S. and Rothberg, J. M., WO9210518-A1 by Yale University. Other studies reporting on the biological functions of proteins having leucine-rich repeats include: Tayar, N., et al., Mol. Cell Endocrinol., (Ireland), 125(1-2):65-70 (Dec. 1996) (gonadotropin receptor involvement); Miura, Y., et al., Nippon Rinsho (Japan), 54(7):1784-1789 (July 1996) (apoptosis involvement); Harris, P. C., et al., J. Am. Soc. Nephrol., 6(4):1125-1133 (Oct. 1995) (kidney disease involvement); and Ruoslahti, E. I., et al., WO9110727-A by La Jolla Cancer Research Foundation (decorin binding to transforming growth factor- β involvement for treatment for cancer, wound healing and scarring). Also of particular interest is fibromodulin and its use to prevent or reduce dermal scarring. A study of fibromodulin is found in U.S. Patent No. 5,654,270 to Ruoslahti, et al.

Efforts are therefore being undertaken by both industry and academia to identify new proteins having leucine rich repeats to better understand protein-protein interactions. Of particular interest are those proteins having leucine rich repeats and homology to known proteins having leucine rich repeats such as fibromodulin, the SLIT protein and platelet glycoprotein V. Many efforts are focused on the screening of mammalian recombinant DNA libraries to identify the coding sequences for novel secreted and membrane-bound proteins having leucine rich repeats. We herein describe the identification and characterization of novel polypeptides having homology to fibromodulin, herein designated as PRO265 polypeptides.

6. PRO219

Human matrilin-2 polypeptide is a member of the von Willebrand factor type A-like module superfamily. von Willebrand factor is a protein which plays an important role in the maintenance of hemostasis. More specifically, von Willebrand factor is a protein which is known to participate in platelet-vessel wall

interactions at the site of vascular injury via its ability to interact and form a complex with Factor VIII. The absence of von Willebrand factor in the blood causes an abnormality with the blood platelets that prevents platelet adhesion to the vascular wall at the site of the vascular injury. The result is the propensity for bruising, nose bleeds, intestinal bleeding, and the like comprising von Willebrand's disease.

Given the physiological importance of the blood clotting factors, efforts are currently being undertaken
5 by both industry and academia to identify new, native proteins which may be involved in the coagulation process. We herein describe the identification of a novel full-length polypeptide which possesses homology to the human matrilin-2 precursor polypeptide.

7. PRO246

10 The cell surface protein HCAR is a membrane-bound protein that acts as a receptor for subgroup C of the adenoviruses and subgroup B of the coxsackieviruses. Thus, HCAR may provide a means for mediating viral infection of cells in that the presence of the HCAR receptor on the cellular surface provides a binding site for viral particles, thereby facilitating viral infection.

In light of the physiological importance of membrane-bound proteins and specifically those which serve
15 a cell surface receptor for viruses, efforts are currently being undertaken by both industry and academia to identify new, native membrane-bound receptor proteins. Many of these efforts are focused on the screening of mammalian recombinant DNA libraries to identify the coding sequences for novel receptor proteins. We herein describe a novel membrane-bound polypeptide (designated herein as PRO246) having homology to the cell surface protein HCAR and to various tumor antigens including A33 and carcinoembryonic antigen, wherein this
20 polypeptide may be a novel cell surface virus receptor or tumor antigen.

8. PRO228

There are a number of known seven transmembrane proteins and within this family is a group which includes CD97 and EMR1. CD97 is a seven-span transmembrane receptor which has a cellular ligand, CD55,
25 DAF. Hamann, et al., J. Exp. Med. (U.S.), 184(3):1189 (1996). Additionally, CD97 has been reported as being a dedifferentiation marker in human thyroid carcinomas and as associated with inflammation. Aust, et al., Cancer Res. (U.S.), 57(9):1798 (1997); Gray, et al., J. Immunol. (U.S.), 157(12):5438 (1996). CD97 has also been reported as being related to the secretin receptor superfamily, but unlike known members of that family, CD97 and EMR1 have extended extracellular regions that possess several EGF domains at the N-terminus.
30 Hamann, et al., Genomics, 32(1):144 (1996); Harmann, et al., J. Immunol., 155(4):1942 (1995). EMR1 is further described in Lin, et al., Genomics, 41(3):301 (1997) and Baud, et al., Genomics, 26(2):334 (1995). While CD97 and EMR1 appear to be related to the secretin receptors, a known member of the secretin family of G protein-coupled receptors includes the alpha-latrotoxin receptor, latrophilin, which has been described as calcium independent and abundant among neuronal tissues. Lelianova, et al., J. Biol. Chem., 272(34), 21504
35 (1997); Davletov, et al., J. Biol. Chem. (U.S.), 271(38):23239 (1996). Both members of the secretin receptor superfamily and non-members which are related to the secretin receptor superfamily, or CRF and calcitonin receptors are of interest. In particular, new members of these families, identified by their homology to known

proteins, are of interest.

Efforts are being undertaken by both industry and academia to identify new membrane-bound receptor proteins, particularly transmembrane proteins with EGF repeats and large N-terminuses which may belong to the family of seven-transmembrane proteins of which CD97 and EMR1 are members. We herein describe the identification and characterization of novel polypeptides having homology to CD97 and EMR1, designated herein 5 as PRO228 polypeptides.

9. PRO533

Growth factors are molecular signals or mediators that enhance cell growth or proliferation, alone or in concert, by binding to specific cell surface receptors. However, there are other cellular reactions than only 10 growth upon expression to growth factors. As a result, growth factors are better characterized as multifunctional and potent cellular regulators. Their biological effects include proliferation, chemotaxis and stimulation of extracellular matrix production. Growth factors can have both stimulatory and inhibitory effects. For example, transforming growth factors (TGF- β) is highly pleiotropic and can stimulate proliferation in some cells, especially connective tissues, while being a potent inhibitor of proliferation in others, such as lymphocytes and 15 epithelial cells.

The physiological effect of growth stimulation or inhibition by growth factors depends upon the state of development and differentiation of the target tissue. The mechanism of local cellular regulation by classical endocrine molecules comprehends autocrine (same cell), juxtacrine (neighbor cell), and paracrine (adjacent cell) pathways. Peptide growth factors are elements of a complex biological language, providing the basis for 20 intercellular communication. They permit cells to convey information between each other, mediate interaction between cells and change gene expression. The effect of these multifunctional and pluripotent factors is dependent on the presence or absence of other peptides.

Fibroblast growth factors (FGFs) are a family of heparin-binding, potent mitogens for both normal diploid fibroblasts and established cell lines, Godpodarowicz, D. et al. (1984), Proc. Natl. Acad. Sci. USA 81: 25 6983. The FGF family comprises acidic FGF (FGF-1), basic FGF (FGF-2), INT-2 (FGF-3), K-FGF/HST (FGF-4), FGF-5, FGF-6, KGF (FGF-7), AIGF (FGF-8) among others. All FGFs have two conserved cysteine residues and share 30-50% sequence homology at the amino acid level. These factors are mitogenic for a wide variety of normal diploid mesoderm-derived and neural crest-derived cells, inducing granulosa cells, adrenal cortical cells, chondrocytes, myoblasts, corneal and vascular endothelial cells (bovine or human), vascular 30 smooth muscle cells, lens, retina and prostatic epithelial cells, oligodendrocytes, astrocytes, chondrocytes, myoblasts and osteoblasts.

Fibroblast growth factors can also stimulate a large number of cell types in a non-mitogenic manner. These activities include promotion of cell migration into a wound area (chemotaxis), initiation of new blood vessel formulation (angiogenesis), modulation of nerve regeneration and survival (neurotrophism), modulation 35 of endocrine functions, and stimulation or suppression of specific cellular protein expression, extracellular matrix production and cell survival. Baird, A. & Bohlen, P., *Handbook of Exp. Pharmacol.* 95(1): 369-418 (1990). These properties provide a basis for using fibroblast growth factors in therapeutic approaches to accelerate

wound healing, nerve repair, collateral blood vessel formation, and the like. For example, fibroblast growth factors, have been suggested to minimize myocardium damage in heart disease and surgery (U.S.P. 4,378,437).

We herein describe the identification and characterization of novel polypeptides having homology to FGF, herein designated PRO533 polypeptides.

5 10. PRO245

Some of the most important proteins involved in the above described regulation and modulation of cellular processes are the enzymes which regulate levels of protein phosphorylation in the cell. For example, it is known that the transduction of signals that regulate cell growth and differentiation is regulated at least in part by phosphorylation and dephosphorylation of various cellular proteins. The enzymes that catalyze these 10 processes include the protein kinases, which function to phosphorylate various cellular proteins, and the protein phosphatases, which function to remove phosphate residues from various cellular proteins. The balance of the level of protein phosphorylation in the cell is thus mediated by the relative activities of these two types of enzymes.

Although many protein kinase enzymes have been identified, the physiological role played by many of 15 these catalytic proteins has yet to be elucidated. It is well known, however, that a number of the known protein kinases function to phosphorylate tyrosine residues in proteins, thereby leading to a variety of different effects. Perhaps most importantly, there has been a great deal of interest in the protein tyrosine kinases since the discovery that many oncogene products and growth factors possess intrinsic protein tyrosine kinase activity. There is, therefore, a desire to identify new members of the protein tyrosine kinase family.

20 Given the physiological importance of the protein kinases, efforts are being undertaken by both industry and academia to identify new, native kinase proteins. Many of these efforts are focused on the screening of mammalian recombinant DNA libraries to identify the coding sequences for novel kinase proteins. We herein describe the identification and characterization of novel polypeptides having homology to tyrosine kinase proteins, designated herein as PRO245 polypeptides.

25 11. PRO220, PRO221 and PRO227

Protein-protein interactions include receptor and antigen complexes and signaling mechanisms. As more is known about the structural and functional mechanisms underlying protein-protein interactions, protein-protein interactions can be more easily manipulated to regulate the particular result of the protein-protein interaction. 30 Thus, the underlying mechanisms of protein-protein interactions are of interest to the scientific and medical community.

All proteins containing leucine-rich repeats are thought to be involved in protein-protein interactions. Leucine-rich repeats are short sequence motifs present in a number of proteins with diverse functions and cellular locations. The crystal structure of ribonuclease inhibitor protein has revealed that leucine-rich repeats 35 correspond to beta-alpha structural units. These units are arranged so that they form a parallel beta-sheet with one surface exposed to solvent, so that the protein acquires an unusual, nonglobular shape. These two features have been indicated as responsible for the protein-binding functions of proteins containing leucine-rich repeats.

See, Kobe and Deisenhofer, Trends Biochem. Sci., 19(10):415-421 (Oct. 1994).

A study has been reported on leucine-rich proteoglycans which serve as tissue organizers, orienting and ordering collagen fibrils during ontogeny and are involved in pathological processes such as wound healing, tissue repair, and tumor stroma formation. Iozzo, R. V., Crit. Rev. Biochem. Mol. Biol., 32(2):141-174 (1997). Other studies implicating leucine rich proteins in wound healing and tissue repair are De La Salle, C., et al., Vouv. Rev. Fr. Hematol. (Germany), 37(4):215-222 (1995), reporting mutations in the leucine rich motif in a complex associated with the bleeding disorder Bernard-Soulier syndrome and Chlemetson, K. J., Thromb. Haemost. (Germany), 74(1):111-116 (July 1995), reporting that platelets have leucine rich repeats. Another protein of particular interest which has been reported to have leucine-rich repeats is the SLIT protein which has been reported to be useful in treating neuro-degenerative diseases such as Alzheimer's disease, nerve damage such as in Parkinson's disease, and for diagnosis of cancer, see, Artavanistsakonas, S. and Rothberg, J. M., WO9210518-A1 by Yale University. Other studies reporting on the biological functions of proteins having leucine-rich repeats include: Tayar, N., et al., Mol. Cell Endocrinol., (Ireland), 125(1-2):65-70 (Dec. 1996) (gonadotropin receptor involvement); Miura, Y., et al., Nippon Rinsho (Japan), 54(7):1784-1789 (July 1996) (apoptosis involvement); Harris, P. C., et al., J. Am. Soc. Nephrol., 6(4):1125-1133 (Oct. 1995) (kidney disease involvement); and Ruoslahti, E. I., et al., WO9110727-A by La Jolla Cancer Research Foundation (decorin binding to transforming growth factor β involvement for treatment for cancer, wound healing and scarring).

Efforts are therefore being undertaken by both industry and academia to identify new proteins having leucine rich repeats to better understand protein-protein interactions. Of particular interest are those proteins having leucine rich repeats and homology to known proteins having leucine rich repeats such as the SLIT protein and platelet glycoprotein V.

12. PRO258

Immunoglobulins are antibody molecules, the proteins that function both as receptors for antigen on the B-cell membrane and as the secreted products of the plasma cell. Like all antibody molecules, immunoglobulins perform two major functions: they bind specifically to an antigen and they participate in a limited number of biological effector functions. Therefore, new members of the Ig superfamily are always of interest. Molecules which act as receptors by various viruses and those which act to regulate immune function are of particular interest. Also of particular interest are those molecules which have homology to known Ig family members which act as virus receptors or regulate immune function. Thus, molecules having homology to poliovirus receptors, CRTAM and CD166 (a ligand for lymphocyte antigen CD6) are of particular interest.

Extracellular and membrane-bound proteins play important roles in the formation, differentiation and maintenance of multicellular organisms. The fate of many individual cells, e.g., proliferation, migration, differentiation, or interaction with other cells, is typically governed by information received from other cells and/or the immediate environment. This information is often transmitted by secreted polypeptides (for instance, mitogenic factors, survival factors, cytotoxic factors, differentiation factors, neuropeptides, and hormones) which are, in turn, received and interpreted by diverse cell receptors or membrane-bound proteins. These secreted

polypeptides or signaling molecules normally pass through the cellular secretory pathway to reach their site of action in the extracellular environment, usually at a membrane-bound receptor protein.

We herein describe the identification and characterization of novel polypeptides having homology to CRTAM, designated herein as PRO258 polypeptides.

5 13. PRO266

Protein-protein interactions include receptor and antigen complexes and signaling mechanisms. As more is known about the structural and functional mechanisms underlying protein-protein interactions, protein-protein interactions can be more easily manipulated to regulate the particular result of the protein-protein interaction. Thus, the underlying mechanisms of protein-protein interactions are of interest to the scientific and medical

10 community.

All proteins containing leucine-rich repeats are thought to be involved in protein-protein interactions. Leucine-rich repeats are short sequence motifs present in a number of proteins with diverse functions and cellular locations. The crystal structure of ribonuclease inhibitor protein has revealed that leucine-rich repeats correspond to beta-alpha structural units. These units are arranged so that they form a parallel beta-sheet with 15 one surface exposed to solvent, so that the protein acquires an unusual, nonglobular shape. These two features have been indicated as responsible for the protein-binding functions of proteins containing leucine-rich repeats. See, Kobe and Deisenhofer, Trends Biochem. Sci., 19(10):415-421 (Oct. 1994).

A study has been reported on leucine-rich proteoglycans which serve as tissue organizers, orienting and ordering collagen fibrils during ontogeny and are involved in pathological processes such as wound healing, 20 tissue repair, and tumor stroma formation. Iozzo, R. V., Crit. Rev. Biochem. Mol. Biol., 32(2):141-174 (1997). Others studies implicating leucine rich proteins in wound healing and tissue repair are De La Salle, C., et al., Vouv. Rev. Fr. Hematol. (Germany), 37(4):215-222 (1995), reporting mutations in the leucine rich motif in a complex associated with the bleeding disorder Bernard-Soulier syndrome and Chlemetson, K. J., Thromb. Haemost. (Germany), 74(1):111-116 (July 1995), reporting that platelets have leucine rich repeats. Another 25 protein of particular interest which has been reported to have leucine-rich repeats is the SLIT protein which has been reported to be useful in treating neuro-degenerative diseases such as Alzheimer's disease, nerve damage such as in Parkinson's disease, and for diagnosis of cancer, see, Artavanistsakonas, S. and Rothberg, J. M., WO9210518-A1 by Yale University. Other studies reporting on the biological functions of proteins having leucine-rich repeats include: Tayar, N., et al., Mol. Cell Endocrinol., (Ireland), 125(1-2):65-70 (Dec. 1996) 30 (gonadotropin receptor involvement); Miura, Y., et al., Nippon Rinsho (Japan), 54(7):1784-1789 (July 1996) (apoptosis involvement); Harris, P. C., et al., J. Am. Soc. Nephrol., 6(4):1125-1133 (Oct. 1995) (kidney disease involvement); and Ruoslahti, E. I., et al., WO9110727-A by La Jolla Cancer Research Foundation (decorin binding to transforming growth factor β involvement for treatment for cancer, wound healing and scarring).

35 Efforts are therefore being undertaken by both industry and academia to identify new proteins having leucine rich repeats to better understand protein-protein interactions, neuronal development and adhesin molecules. Of particular interest are those proteins having leucine rich repeats and homology to known proteins

having leucine rich repeats such as the SLIT protein. We herein describe novel polypeptides having homology to SLIT, designated herein as PRO266 polypeptides.

14. **PRO269**

Thrombomodulin binds to and regulates the activity of thrombin. It is important in the control of blood coagulation. Thrombomodulin functions as a natural anticoagulant by accelerating the activation of protein C by thrombin. Soluble thrombomodulin may have therapeutic use as an anti-thrombotic agent with reduced risk for hemorrhage as compared with heparin. Thrombomodulin is a cell surface trans-membrane glycoprotein, present on endothelial cells and platelets. A smaller, functionally active form of thrombomodulin circulates in the plasma and is also found in urine. (In Haeberli, A., Human Protein Data, VCH Oub., N.Y., 1992). Peptides having homology to thrombomodulin are particularly desirable.

We herein describe the identification and characterization of novel polypeptides having homology to thrombomodulin, designated herein as PRO269 polypeptides.

15. **PRO287**

15 Procollagen C-proteinase enhancer protein binds to and enhances the activity of bone morphogenic protein "BMP1"/procollagen C-proteinase (PCP). It plays a role in extracellular matrix deposition. BMP1 proteins may be used to induce bone and/or cartilage formation and in wound healing and tissue repair. Therefore, procollagen C-proteinase enhancer protein, BMP1 and proteins having homology thereto, are of interest to the scientific and medical communities.

20 We herein describe the identification and characterization of novel polypeptides having homology to procollagen C-proteinase enhancer protein precursor and procollagen C-proteinase enhancer protein, designated herein as PRO287 polypeptides.

16. **PRO214**

25 Growth factors are molecular signals or mediators that enhances cell growth or proliferation, alone or in concert, by binding to specific cell surface receptors. However, there are other cellular reactions than only growth upon expression to growth factors. As a result, growth factors are better characterized as multifunctional and potent cellular regulators. Their biological effects include proliferation, chemotaxis and stimulation of extracellular matrix production. Growth factors can have both stimulatory and inhibitory effects. For example, 30 transforming growth factor β (TGF- β) is highly pleiotropic and can stimulate proliferation in some cells, especially connective tissue, while being a potent inhibitor of proliferation in others, such as lymphocytes and epithelial cells.

The physiological effect of growth stimulation or inhibition by growth factors depends upon the state of development and differentiation of the target tissue. The mechanism of local cellular regulation by classical 35 endocrine molecules involves comprehends autocrine (same cell), juxtacrine (neighbor cell), and paracrine (adjacent cells) pathways. Peptide growth factors are elements of a complex biological language, providing the basis for intercellular communication. They permit cells to convey information between each other, mediate

interaction between cells and change gene expression. The effect of these multifunctional and pluripotent factors is dependent on the presence or absence of other peptides.

Epidermal growth factor (EGF) is a conventional mitogenic factor that stimulates the proliferation of various types of cells including epithelial cells and fibroblasts. EGF binds to and activates the EGF receptor (EGFR), which initiates intracellular signaling and subsequent effects. The EGFR is expressed in neurons of the cerebral cortex, cerebellum, and hippocampus in addition to other regions of the central nervous system (CNS). In addition, EGF is also expressed in various regions of the CNS. Therefore, EGF acts not only on mitotic cells, but also on postmitotic neurons. In fact, many studies have indicated that EGF has neurotrophic or neuromodulatory effects on various types of neurons in the CNS. For example, EGF acts directly on cultured cerebral cortical and cerebellar neurons, enhancing neurite outgrowth and survival. On the other hand, EGF also acts on other cell types, including septal cholinergic and mesencephalic dopaminergic neurons, indirectly through glial cells. Evidence of the effects of EGF on neurons in the CNS is accumulating, but the mechanisms of action remain essentially unknown. EGF-induced signaling in mitotic cells is better understood than in postmitotic neurons. Studies of cloned pheochromocytoma PC12 cells and cultured cerebral cortical neurons have suggested that the EGF-induced neurotrophic actions are mediated by sustained activation of the EGFR and mitogen-activated protein kinase (MAPK) in response to EGF. The sustained intracellular signaling correlates with the decreased rate of EGFR down-regulation, which might determine the response of neuronal cells to EGF. It is likely that EGF is a multi-potent growth factor that acts upon various types of cells including mitotic cells and postmitotic neurons.

EGF is produced by the salivary and Brunner's glands of the gastrointestinal system, kidney, pancreas, thyroid gland, pituitary gland, and the nervous system, and is found in body fluids such as saliva, blood, cerebrospinal fluid (CSF), urine, amniotic fluid, prostatic fluid, pancreatic juice, and breast milk, Plata-Salaman, *CR Peptides* 12: 653-663 (1991).

EGF is mediated by its membrane specific receptor, which contains an intrinsic tyrosine kinase. Stoscheck CM *et al.*, *J. Cell Biochem.* 31: 135-152 (1986). EGF is believed to function by binding to the extracellular portion of its receptor which induces a transmembrane signal that activates the intrinsic tyrosine kinase.

Purification and sequence analysis of the EGF-like domain has revealed the presence of six conserved cysteine residues which cross-bind to create three peptide loops, Savage *CR* *et al.*, *J. Biol. Chem.* 248: 7669-7672 (1979). It is now generally known that several other peptides can react with the EGF receptor which share the same generalized motif $X_nCX_2CX_{4/5}CX_{10}CXCX_5GX_2CX_n$, where X represents any non-cysteine amino acid, and n is a variable repeat number. Non isolated peptides having this motif include TGF- α , amphiregulin, schwannoma-derived growth factor (SDGF), heparin-binding EGF-like growth factors and certain virally encoded peptides (e.g., Vaccinia virus, Reisner AH, *Nature* 313: 801-803 (1985), Shope fibroma virus, Chang W., *et al.*, *Mol Cell Biol.* 7: 535-540 (1987), Molluscum contagiosum, Porter CD & Archard LC, *J. Gen. Virol.* 68: 673-682 (1987), and Myxoma virus, Upton C *et al.*, *J. Virol.* 61: 1271-1275 (1987). Prigent SA & Lemoine N.R., *Prog. Growth Factor Res.* 4: 1-24 (1992).

EGF-like domains are not confined to growth factors but have been observed in a variety of cell-surface and extracellular proteins which have interesting properties in cell adhesion, protein-protein interaction and development, Laurence DJR & Gusterson BA, *Tumor Biol.* 11: 229-261 (1990). These proteins include blood coagulation factors (factors VI, IX, X, XII, protein C, protein S, protein Z, tissue plasminogen activator, urokinase), extracellular matrix components (laminin, cytотactин, entactin), cell surface receptors (LDL receptor, 5 thrombomodulin receptor) and immunity-related proteins (complement C1r, uromodulin).

Even more interesting, the general structure pattern of EGF-like precursors is preserved through lower organisms as well as in mammalian cells. A number of genes with developmental significance have been identified in invertebrates with EGF-like repeats. For example, the *notch* gene of *Drosophila* encodes 36 tandemly arranged 40 amino acid repeats which show homology to EGF, Wharton W *et al.*, *Cell* 43: 557-581 10 (1985). Hydropathy plots indicate a putative membrane spanning domain, with the EGF-related sequences being located on the extracellular side of the membrane. Other homeotic genes with EGF-like repeats include Delta, 95F and 5ZD which were identified using probes based on Notch, and the nematode gene *Lin-12* which encodes a putative receptor for a developmental signal transmitted between two specified cells.

Specifically, EGF has been shown to have potential in the preservation and maintenance of 15 gastrointestinal mucosa and the repair of acute and chronic mucosal lesions, Konturek, *PC* *et al.*, *Eur. J. Gastroenterol Hepatol.* 7 (10), 933-37 (1995), including the treatment of necrotizing enterocolitis, Zollinger-Ellison syndrome, gastrointestinal ulceration gastrointestinal ulcerations and congenital microvillus atrophy, A. Guglietta & PB Sullivan, *Eur. J. Gastroenterol Hepatol.* 7(10), 945-50 (1995). Additionally, EGF has been implicated in hair follicle differentiation; C.L. du Cros, *J. Invest. Dermatol.* 101 (1 Suppl.), 106S-113S (1993), 20 SG Hillier, *Clin. Endocrinol.* 33(4), 427-28 (1990); kidney function, L.L. Hamm *et al.*, *Semin. Nephrol.* 13(1): 109-15 (1993), RC Harris, *Am. J. Kidney Dis.* 17(6): 627-30 (1991); tear fluid, GB van Setten *et al.*, *Int. Ophthalmol* 15(6): 359-62 (1991); vitamin K mediated blood coagulation, J. Stenflo *et al.*, *Blood* 78(7): 1637-51 (1991). EGF is also implicated various skin disease characterized by abnormal keratinocyte differentiation, e.g., psoriasis, epithelial cancers such as squamous cell carcinomas of the lung, epidermoid carcinoma of the vulva 25 and gliomas. King, LE *et al.*, *Am. J. Med. Sci.* 296: 154-158 (1988).

Of great interest is mounting evidence that genetic alterations in growth factors signaling pathways are closely linked to developmental abnormalities and to chronic diseases including cancer. Aaronson SA, *Science* 254: 1146-1153 (1991). For example, c-erb-2 (also known as HER-2), a proto-oncogene with close structural similarity to EGF receptor protein, is overexpressed in human breast cancer. King *et al.*, *Science* 229: 974-976 30 (1985); Gullick, WJ, *Hormones and their actions*, Cooke BA *et al.*, eds, Amsterdam, Elsevier, pp 349-360 (1986).

17. PRO317

The TGF- β supergene family, or simply TGF- β superfamily, a group of secreted proteins, includes 35 a large number of related growth and differentiation factors expressed in virtually all phyla. Superfamily members bind to specific cell surface receptors that activate signal transduction mechanisms to elicit their multifunctional cytokine effects. Kolodziejczyk and Hall, *Biochem. Cell. Biol.*, 74: 299-314 (1996); Attisano

and Wrana, Cytokine Growth Factor Rev., **7**: 327-339 (1996); and Hill, Cellular Signaling, **8**: 533-544 (1996).

Members of this family include five distinct forms of TGF- β (Sporn and Roberts, in Peptide Growth Factors and Their Receptors, Sporn and Roberts, eds. (Springer-Verlag: Berlin, 1990) pp. 419-472), as well as the differentiation factors vgl (Weeks and Melton, Cell, **51**: 861-867 (1987)) and DPP-C polypeptide (Padgett *et al.*, Nature, **325**: 81-84 (1987)), the hormones activin and inhibin (Mason *et al.*, Nature, **318**: 659-663 (1985); 5 Mason *et al.*, Growth Factors, **1**: 77-88 (1987)), the Mullerian-inhibiting substance (MIS) (Cate *et al.*, Cell, **45**: 685-698 (1986)), the bone morphogenetic proteins (BMPs) (Wozney *et al.*, Science, **242**: 1528-1534 (1988); PCT WO 88/00205 published January 14, 1988; U.S. 4,877,864 issued October 31, 1989), the developmentally regulated proteins Vgr-1 (Lyons *et al.*, Proc. Natl. Acad. Sci. USA, **86**: 4554-4558 (1989)) and Vgr-2 (Jones *et al.*, Molec. Endocrinol., **6**: 1961-1968 (1992)), the mouse growth differentiation factor (GDF), such as GDF- 10 3 and GDF-9 (Kingsley, Genes Dev., **8**: 133-146 (1994); McPherron and Lee, J. Biol. Chem., **268**: 3444-3449 (1993)), the mouse lefty/Stra1 (Meno *et al.*, Nature, **381**: 151-155 (1996); Bouillet *et al.*, Dev. Biol., **170**: 420-433 (1995)), glial cell line-derived neurotrophic factor (GDNF) (Lin *et al.*, Science, **260**: 1130-1132 (1993)), neuritin (Kotzbauer *et al.*, Nature, **384**: 467-470 (1996)), and endometrial bleeding-associated factor (EBAF) (Kothapalli *et al.*, J. Clin. Invest., **99**: 2342-2350 (1997)). The subset BMP-2A and BMP-2B is approximately 15 75% homologous in sequence to DPP-C and may represent the mammalian equivalent of that protein.

The proteins of the TGF- β superfamily are disulfide-linked homo- or heterodimers encoded by larger precursor polypeptide chains containing a hydrophobic signal sequence, a long and relatively poorly conserved N-terminal pro region of several hundred amino acids, a cleavage site (usually polybasic), and a shorter and more highly conserved C-terminal region. This C-terminal region corresponds to the processed mature protein 20 and contains approximately 100 amino acids with a characteristic cysteine motif, *i.e.*, the conservation of seven of the nine cysteine residues of TGF- β among all known family members. Although the position of the cleavage site between the mature and pro regions varies among the family members, the C-terminus of all of the proteins is in the identical position, ending in the sequence Cys-X-Cys-X, but differing in every case from the TGF- β consensus C-terminus of Cys-Lys-Cys-Ser. Sporn and Roberts, 1990, *supra*.

25 There are at least five forms of TGF- β currently identified, TGF- β 1, TGF- β 2, TGF- β 3, TGF- β 4, and TGF- β 5. The activated form of TGF- β 1 is a homodimer formed by dimerization of the carboxy-terminal 112 amino acids of a 390 amino acid precursor. Recombinant TGF- β 1 has been cloned (Derynck *et al.*, Nature, **316**: 701-705 (1985)) and expressed in Chinese hamster ovary cells (Gentry *et al.*, Mol. Cell. Biol., **7**: 3418-3427 (1987)). Additionally, recombinant human TGF- β 2 (deMartin *et al.*, EMBO J., **6**: 3673 (1987)), as well as 30 human and porcine TGF- β 3 (Derynck *et al.*, EMBO J., **7**: 3737-3743 (1988); ten Dijke *et al.*, Proc. Natl. Acad. Sci. USA, **85**: 4715 (1988)) have been cloned. TGF- β 2 has a precursor form of 414 amino acids and is also processed to a homodimer from the carboxy-terminal 112 amino acids that shares approximately 70% homology with the active form of TGF- β 1 (Marquardt *et al.*, J. Biol. Chem., **262**: 12127 (1987)). See also EP 200,341; 169,016; 268,561; and 267,463; U.S. Pat. No. 4,774,322; Cheifetz *et al.*, Cell, **48**: 409-415 (1987); Jakowlew 35 *et al.*, Molecular Endocrin., **2**: 747-755 (1988); Derynck *et al.*, J. Biol. Chem., **261**: 4377-4379 (1986); Sharples *et al.*, DNA, **6**: 239-244 (1987); Derynck *et al.*, Nucl. Acids. Res., **15**: 3188-3189 (1987); Derynck *et al.*, Nucl. Acids. Res., **15**: 3187 (1987); Seyedin *et al.*, J. Biol. Chem., **261**: 5693-5695 (1986); Madisen *et*

al., DNA, 7: 1-8 (1988); and Hanks *et al.*, Proc. Natl. Acad. Sci. (U.S.A.), 85: 79-82 (1988).

TGF- β 4 and TGF- β 5 were cloned from a chicken chondrocyte cDNA library (Jakowlew *et al.*, Molec. Endocrinol., 2: 1186-1195 (1988)) and from a frog oocyte cDNA library, respectively.

The pro region of TGF- β associates non-covalently with the mature TGF- β dimer (Wakefield *et al.*, J. Biol. Chem., 263: 7646-7654 (1988); Wakefield *et al.*, Growth Factors, 1: 203-218 (1989)), and the pro regions are found to be necessary for proper folding and secretion of the active mature dimers of both TGF- β and activin (Gray and Mason, Science, 247: 1328-1330 (1990)). The association between the mature and pro regions of TGF- β masks the biological activity of the mature dimer, resulting in formation of an inactive latent form. Latency is not a constant of the TGF- β superfamily, since the presence of the pro region has no effect on activin or inhibin biological activity.

10 A unifying feature of the biology of the proteins from the TGF- β superfamily is their ability to regulate developmental processes. TGF- β has been shown to have numerous regulatory actions on a wide variety of both normal and neoplastic cells. TGF- β is multifunctional, as it can either stimulate or inhibit cell proliferation, differentiation, and other critical processes in cell function (Sporn and Roberts, *supra*).

15 One member of the TGF- β superfamily, EBAF, is expressed in endometrium only in the late secretory phase and during abnormal endometrial bleeding. Kothapalli *et al.*, J. Clin. Invest., 99: 2342-2350 (1997). Human endometrium is unique in that it is the only tissue in the body that bleeds at regular intervals. In addition, abnormal endometrial bleeding is one of the most common manifestations of gynecological diseases, and is a prime indication for hysterectomy. *In situ* hybridization showed that the mRNA of EBAF was expressed in the stroma without any significant mRNA expression in the endometrial glands or endothelial cells.

20 The predicted protein sequence of EBAF showed a strong homology to the protein encoded by mouse *lefty/stra3* of the TGF- β superfamily. A motif search revealed that the predicted EBAF protein contains most of the cysteine residues which are conserved among the TGF- β -related proteins and which are necessary for the formation of the cysteine knot structure. The EBAF sequence contains an additional cysteine residue, 12 amino acids upstream from the first conserved cysteine residue. The only other family members known to contain an additional cysteine residue are TGF- β s, inhibins, and GDF-3. EBAF, similar to LEFTY, GDF-3/Vgr2, and GDF-9, lacks the cysteine residue that is known to form the intermolecular disulfide bond. Therefore, EBAF appears to be an additional member of the TGF- β superfamily with an unpaired cysteine residue that may not exist as a dimer. However, hydrophobic contacts between the two monomer subunits may promote dimer formation. Fluorescence *in situ* hybridization showed that the *ebaf* gene is located on human chromosome 1 at 25 band q42.1.

30 Additional members of the TGF- β superfamily, such as those related to EBAF, are being searched for by industry and academics. We herein describe the identification and characterization of novel polypeptides having homology to EBAF, designated herein as PRO317 polypeptides.

35 18. PRO301

The widespread occurrence of cancer has prompted the devotion of considerable resources and discovering new treatments of treatment. One particular method involves the creation of tumor or cancer specific

monoclonal antibodies (mAbs) which are specific to tumor antigens. Such mAbs, which can distinguish between normal and cancerous cells are useful in the diagnosis, prognosis and treatment of the disease. Particular antigens are known to be associated with neoplastic diseases, such as colorectal cancer.

One particular antigen, the A33 antigen is expressed in more than 90% of primary or metastatic colon cancers as well as normal colon epithelium. Since colon cancer is a widespread disease, early diagnosis and treatment is an important medical goal. Diagnosis and treatment of colon cancer can be implemented using 5 monoclonal antibodies (mAbs) specific therefore having fluorescent, nuclear magnetic or radioactive tags. Radioactive gene, toxins and/or drug tagged mAbs can be used for treatment *in situ* with minimal patient description. mAbs can also be used to diagnose during the diagnosis and treatment of colon cancers. For example, when the serum levels of the A33 antigen are elevated in a patient, a drop of the levels after surgery 10 would indicate the tumor resection was successful. On the other hand, a subsequent rise in serum A33 antigen levels after surgery would indicate that metastases of the original tumor may have formed or that new primary tumors may have appeared. Such monoclonal antibodies can be used in lieu of, or in conjunction with surgery and/or other chemotherapies. For example, U.S.P. 4,579,827 and U.S.S.N. 424,991 (E.P. 199,141) are directed to therapeutic administration of monoclonal antibodies, the latter of which relates to the application of 15 anti-A33 mAb.

Many cancers of epithelial origin have adenovirus receptors. In fact, adenovirus-derived vectors have been proposed as a means of inserting antisense nucleic acids into tumors (U.S.P. 5,518,885). Thus, the association of viral receptors with neoplastic tumors is not unexpected.

We herein describe the identification and characterization of novel polypeptides having homology to 20 certain cancer-associated antigens, designated herein as PRO301 polypeptides.

19. PRO224

Cholesterol uptake can have serious implications on one's health. Cholesterol uptake provides cells with most of the cholesterol they require for membrane synthesis. If this uptake is blocked, cholesterol accumulates 25 in the blood and can contribute to the formation of atherosclerotic plaques in blood vessel walls. Most cholesterol is transported in the blood bound to protein in the form of complexes known as low-density lipoproteins (LDLs). LDLs are endocytosed into cells via LDL receptor proteins. Therefore, LDL receptor proteins, and proteins having homology thereto, are of interest to the scientific and medical communities.

Membrane-bound proteins and receptors can play an important role in the formation, differentiation and 30 maintenance of multicellular organisms. The LDL receptors are an example of membrane-bound proteins which are involved in the synthesis and formation of cell membranes, wherein the health of an individual is affected directly and indirectly by its function. Many membrane-bound proteins act as receptors such as the LDL receptor. These receptors can function to endocytose substrates or they can function as a receptor for a channel. Other membrane-bound proteins function as signals or antigens.

35 Membrane-bound proteins and receptor molecules have various industrial applications, including as pharmaceutical and diagnostic agents. The membrane-bound proteins can also be employed for screening of potential peptide or small molecule regulators of the relevant receptor/ligand interaction. In the case of the LDL

receptor, it is desirable to find molecules which enhance endocytosis so as to lower blood cholesterol levels and plaque formation. It is also desirable to identify molecules which inhibit endocytosis so that these molecules can be avoided or regulated by individuals having high blood cholesterol. Polypeptides which are homologous to lipoprotein receptors but which do not function as lipoprotein receptors are also of interest in the determination of the function of the fragments which show homology.

5 The following studies report on previously known low density lipoprotein receptors and related proteins including apolipoproteins: Sawamura, et al., Nippon Chemiphar Co, Japan patent application J09098787; Novak, S., et al., J. Biol. Chem., 271:(20)11732-6 (1996); Blaas, D., J. Virol., 69(11)7244-7 (Nov. 1995); Scott, J., J. Inherit. Metab. Dis. (UK), 9/Supp. 1 (3-16) (1986); Yamamoto, et al., Cell, 39:27-38 (1984); Rebec, et al., Neurobiol. Aging, 15:5117 (1994); Novak, S., et al., J. Biol. Chemistry, 271:11732-11736 (1996); and Sestavel 10 and Fruchart, Cell Mol. Biol., 40(4):461-81 (June 1994). These publications and others published prior to the filing of this application provide further background to peptides already known in the art.

Efforts are being undertaken by both industry and academia to identify new, native membrane-bound receptor proteins, particularly those having homology to lipoprotein receptors. We herein describe the 15 identification and characterization of novel polypeptides having homology to lipoprotein receptors, designated herein as PRO224 polypeptides.

20. PRO222

Complement is a group of proteins found in the blood that are important in humoral immunity and inflammation. Complement proteins are sequentially activated by antigen-antibody complexes or by proteolytic 20 enzymes. When activated, complement proteins kill bacteria and other microorganisms, affect vascular permeability, release histamine and attract white blood cells. Complement also enhances phagocytosis when bound to target cells. In order to prevent harm to autologous cells, the complement activation pathway is tightly regulated.

Deficiencies in the regulation of complement activation or in the complement proteins themselves may 25 lead to immune-complex diseases, such as systemic lupus erythematosus, and may result in increased susceptibility to bacterial infection. In all cases, early detection of complement deficiency is desirable so that the patient can begin treatment. Thus, research efforts are currently directed toward identification of soluble and membrane proteins that regulate complement activation.

Proteins known to be important in regulating complement activation in humans include Factor H and 30 Complement receptor type 1 (CR1). Factor H is a 150 kD soluble serum protein that interacts with complement protein C3b to accelerate the decay of C3 convertase and acts as a cofactor for Factor I-mediated cleavage of complement protein C4b. Complement receptor type 1 is a 190-280 kD membrane bound protein found in mast cells and most blood cells. CR1 interacts with complement proteins C3b, C4b, and iC3b to accelerate dissociation of C3 convertases, acts as a cofactor for Factor I-mediated cleavage of C3b and C4b, and binds 35 immune complexes and promotes their dissolution and phagocytosis.

Proteins which have homology to complement proteins are of particular interest to the medical and industrial communities. Often, proteins having homology to each other have similar function. It is also of

interest when proteins having homology do not have similar functions, indicating that certain structural motifs identify information other than function, such as locality of function.

Efforts are being undertaken by both industry and academia to identify new, native secreted and membrane-bound proteins, particularly those having homology to known proteins involved in the complement pathway. Proteins involved in the complement pathway were reviewed in Birmingham DJ (1995), Critical

5 Reviews in Immunology, 15(2):133-154 and in Abbas AK, et al. (1994) Cellular and Molecular Immunology, 2nd Ed. W.B. Saunders Company, Philadelphia, pp 295-315.

We herein describe the identification and characterization of novel polypeptides having homology to complement receptors, designated herein as PRO222 polypeptides.

10 21. **PRO234**

The successful function of many systems within multicellular organisms is dependent on cell-cell interactions. Such interactions are affected by the alignment of particular ligands with particular receptors in a manner which allows for ligand-receptor binding and thus a cell-cell adhesion. While protein-protein interactions in cell recognition have been recognized for some time, only recently has the role of carbohydrates 15 in physiologically relevant recognition been widely considered (see B.K. Brandley *et al.*, *J. Leuk. Biol.* **40**: 97 (1986) and N. Sharon *et al.*, *Science* **246**: 227 (1989)). Oligosaccharides are well positioned to act as recognition novel lectins due to their cell surface location and structural diversity. Many oligosaccharide structures can be created through the differential activities of a smaller number of glycosyltransferases. The diverse structures of oligosaccharides can be generated by transcription of relatively few gene products, which suggests that the 20 oligosaccharides are a plausible mechanism by which is directed a wide range of cell-cell interactions. Examples of differential expression of cell surface carbohydrates and putative carbohydrate binding proteins (lectins) on interacting cells have been described (J. Dodd & T.M. Jessel, *J. Neurosci.* **5**: 3278 (1985); L.J. Regan *et al.*, *Proc. Natl. Acad. Sci. USA* **83**: 2248 (1986); M. Constantine-Paton *et al.*, *Nature* **324**: 459 (1986); and M. Tiemeyer *et al.*, *J. Biol. Chem.* **263**: 1671 (1989)). One interesting member of the lectin family are selectins.

25 The migration of leukocytes to sites of acute or chronic inflammation involves adhesive interactions between these cells and the endothelium. This specific adhesion is the initial event in the cascade that is initiated by inflammatory insults, and it is, therefore, of paramount importance to the regulated defense of the organism.

The types of cell adhesion molecules that are involved in the interaction between leukocytes and the 30 endothelium during an inflammatory response currently stands at four: (1) selectins; (2) (carbohydrate and glycoprotein) ligands for selectins; (3) integrins; and (4) integrin ligands, which are members of the immunoglobulin gene superfamily.

The selectins are cell adhesion molecules that are unified both structurally and functionally. Structurally, selectins are characterized by the inclusion of a domain with homology to a calcium-dependent 35 lectin (C-lectins), an epidermal growth factor (egf)-like domain and several complement binding-like domains, Bevilacqua, M.P. *et al.*, *Science* **243**: 1160-1165 (1989); Johnston *et al.*, *Cell* **56**: 1033-1044 (1989); Lasky *et al.*, *Cell* **56**: 1045-1055 (1989); Siegalman, M. *et al.*, *Science* **243**: 1165-1172 (1989); Stoolman, L.M., *Cell* **56**: 907-910 (1989). Functionally, selectins share the common property of their ability to mediate cell binding

through interactions between their lectin domains and cell surface carbohydrate ligands (Brandley, B., *et al.*, *Cell* **63**, 861-863 (1990); Springer, T. and Lasky, L.A., *Nature* **349**, 19-197 (1991); Bevilacqua, M.P. and Nelson, R.M., *J. Clin. Invest.* **91** 379-387 (1993) and Tedder *et al.*, *J. Exp. Med.* **170**: 123-133 (1989).

There are three members identified so far in the selectin family of cell adhesion molecules: L-selectin (also called peripheral lymph node homing receptor (pnHR), LEC-CAM-1, LAM-1, gp90^{MEL}, gp100^{MEL}, 5 gp110^{MEL}, MEL-14 antigen, Leu-8 antigen, TQ-1 antigen, DREG antigen), E-selectin (LEC-CAM-2, LECAM-2, ELAM-1) and P-selectin (LEC-CAM-3, LECAM-3, GMP-140, PADGEM).

The identification of the C-lectin domain has led to an intense effort to define carbohydrate binding ligands for proteins containing such domains. E-selectin is believed to recognize the carbohydrate sequence NeuNAc₂-3Gal β 1-4(Fuc α 1-3)GlcNAc (sialyl-Lewis x, or sLe^x) and related oligosaccharides, Berg *et al.*, *J. Biol. Chem.* **265**: 14869-14872 (1991); Lowe *et al.*, *Cell* **63**: 475-484 (1990); Phillips *et al.*, *Science* **250**: 1130-10 1132 (1990); Tiemeyer *et al.*, *Proc. Natl. Acad. Sci. USA* **88**: 1138-1142 (1991).

L-selectin, which comprises a lectin domain, performs its adhesive function by recognizing carbohydrate-containing ligands on endothelial cells. L-selectin is expressed on the surface of leukocytes, such as lymphocytes, neutrophils, monocytes and eosinophils, and is involved with the trafficking of lymphocytes to 15 peripheral lymphoid tissues (*Gallatin et al.*, *Nature* **303**: 30-34 (1983)) and with acute neutrophil-mediated inflammatory responses (Watson, S.R., *Nature* **349**: 164-167 (1991)). The amino acid sequence of L-selectin and the encoding nucleic acid sequence are, for example, disclosed in U.S. patent No. 5,098,833 issued 24 March 1992.

L-selectin (LECAM-1) is particularly interesting because of its ability to block neutrophil influx (Watson *et al.*, *Nature* **349**: 164-167 (1991)). It is expressed in chronic lymphocytic leukemia cells which bind to HEV (Spertini *et al.*, *Nature* **349**: 691-694 (1991)). It is also believed that HEV structures at sites of chronic inflammation are associated with the symptoms of diseases such as rheumatoid arthritis, psoriasis and multiple sclerosis.

E-selectin (ELAM-1), is particularly interesting because of its transient expression on endothelial cells 25 in response to IL-1 or TNF. Bevilacqua *et al.*, *Science* **243**: 1160 (1989). The time course of this induced expression (2-8 h) suggests a role for this receptor in initial neutrophil induced extravasation in response to infection and injury. It has further been reported that anti-ELAM-1 antibody blocks the influx of neutrophils in a primate asthma model and thus is beneficial for preventing airway obstruction resulting from the inflammatory response. Gundel *et al.*, *J. Clin. Invest.* **88**: 1407 (1991).

The adhesion of circulating neutrophils to stimulated vascular endothelium is a primary event of the 30 inflammatory response. P-selectin has been reported to recognize the Lewis x structure (Gal β 1-4(Fuc α 1-3)GlcNAc), Larsen *et al.*, *Cell* **63**: 467-474 (1990). Others report that an additional terminal linked sialic acid is required for high affinity binding, Moore *et al.*, *J. Cell. Biol.* **112**: 491-499 (1991). P-selectin has been shown to be significant in acute lung injury. Anti-P-selectin antibody has been shown to have strong protective effects in a rodent lung injury model. M.S. Mulligan *et al.*, *J. Clin. Invest.* **90**: 1600 (1991).

We herein describe the identification and characterization of novel polypeptides having homology to lectin proteins, herein designated as PRO234 polypeptides.

22. PRO231

Some of the most important proteins involved in the above described regulation and modulation of cellular processes are the enzymes which regulate levels of protein phosphorylation in the cell. For example, it is known that the transduction of signals that regulate cell growth and differentiation is regulated at least in part by phosphorylation and dephosphorylation of various cellular proteins. The enzymes that catalyze these processes include the protein kinases, which function to phosphorylate various cellular proteins, and the protein phosphatases, which function to remove phosphate residues from various cellular proteins. The balance of the level of protein phosphorylation in the cell is thus mediated by the relative activities of these two types of enzymes.

Protein phosphatases represent a growing family of enzymes that are found in many diverse forms, including both membrane-bound and soluble forms. While many protein phosphatases have been described, the functions of only a very few are beginning to be understood (Tonks, *Semin. Cell Biol.* 4:373-453 (1993) and Dixon, *Recent Prog. Horm. Res.* 51:405-414 (1996)). However, in general, it appears that many of the protein phosphatases function to modulate the positive or negative signals induced by various protein kinases. Therefore, it is likely that protein phosphatases play critical roles in numerous and diverse cellular processes.

Given the physiological importance of the protein phosphatases, efforts are being undertaken by both industry and academia to identify new, native phosphatase proteins. Many of these efforts are focused on the screening of mammalian recombinant DNA libraries to identify the coding sequences for novel phosphatase proteins. Examples of screening methods and techniques are described in the literature [see, for example, Klein et al., *Proc. Natl. Acad. Sci.*, 93:7108-7113 (1996); U.S. Patent No. 5,536,637)].

We herein describe the identification and characterization of novel polypeptides having homology to acid phosphatases, designated herein as PRO231 polypeptides.

23. PRO229

Scavenger receptors are known to protect IgG molecules from catabolic degradation. Riechmann and Hollinger, *Nature Biotechnology*, 15:617 (1997). In particular, studies of the CH2 and CH3 domains have shown that specific sequences of these domains are important in determining the half-lives of antibodies. Ellerson, et al., *J. Immunol.*, 116: 510 (1976); Yasmeen, et al., *J. Immunol.* 116: 518 (1976); Pollock, et al., *Eur. J. Immunol.*, 20: 2021 (1990). Scavenger receptor proteins and antibodies thereto are further reported in U.S. Patent No. 5,510,466 to Krieger, et al. Due to the ability of scavenger receptors to increase the half-life of polypeptides and their involvement in immune function, molecules having homology to scavenger receptors are of importance to the scientific and medical community.

Efforts are being undertaken by both industry and academia to identify new, native secreted and membrane-bound receptor proteins, particularly those having homology to scavenger receptors. Many efforts are focused on the screening of mammalian recombinant DNA libraries to identify the coding sequences for novel secreted and membrane-bound receptor proteins. Examples of screening methods and techniques are described in the literature [see, for example, Klein et al., *Proc. Natl. Acad. Sci.*, 93:7108-7113 (1996); U.S. Patent No. 5,536,637)].

We herein describe the identification and characterization of novel polypeptides having homology to scavenger receptors, designated herein as PRO229 polypeptides.

24. **PRO238**

Oxygen free radicals and antioxidants appear to play an important role in the central nervous system 5 after cerebral ischemia and reperfusion. Moreover, cardiac injury, related to ischaemia and reperfusion has been reported to be caused by the action of free radicals. Additionally, studies have reported that the redox state of the cell is a pivotal determinant of the fate of the cells. Furthermore, reactive oxygen species have been reported to be cytotoxic, causing inflammatory disease, including tissue necrosis, organ failure, atherosclerosis, infertility, birth defects, premature aging, mutations and malignancy. Thus, the control of oxidation and reduction is 10 important for a number of reasons including for control and prevention of strokes, heart attacks, oxidative stress and hypertension. In this regard, reductases, and particularly, oxidoreductases, are of interest. Publications further describing this subject matter include Kelsey, et al., Br. J. Cancer, 76(7):852-4 (1997); Friedrich and Weiss, J. Theor. Biol., 187(4):529-40 (1997) and Pieulle, et al., J. Bacteriol., 179(18):5684-92 (1997).

Efforts are being undertaken by both industry and academia to identify new, native secreted and 15 membrane-bound receptor proteins, particularly secreted proteins which have homology to reductase. Many efforts are focused on the screening of mammalian recombinant DNA libraries to identify the coding sequences for novel secreted and membrane-bound receptor proteins. Examples of screening methods and techniques are described in the literature [see, for example, Klein et al., Proc. Natl. Acad. Sci., 93:7108-7113 (1996); U.S. Patent No. 5,536,637].

20 We herein describe the identification and characterization of novel polypeptides having homology to reductase, designated herein as PRO238 polypeptides.

25. **PRO233**

Studies have reported that the redox state of the cell is an important determinant of the fate of the cell. 25 Furthermore, reactive oxygen species have been reported to be cytotoxic, causing inflammatory disease, including tissue necrosis, organ failure, atherosclerosis, infertility, birth defects, premature aging, mutations and malignancy. Thus, the control of oxidation and reduction is important for a number of reasons, including the control and prevention of strokes, heart attacks, oxidative stress and hypertension. Oxygen free radicals and antioxidants appear to play an important role in the central nervous system after cerebral ischemia and 30 reperfusion. Moreover, cardiac injury, related to ischaemia and reperfusion has been reported to be caused by the action of free radicals. In this regard, reductases, and particularly, oxidoreductases, are of interest. In addition, the transcription factors, NF-kappa B and AP-1, are known to be regulated by redox state and to affect the expression of a large variety of genes thought to be involved in the pathogenesis of AIDS, cancer, atherosclerosis and diabetic complications. Publications further describing this subject matter include Kelsey, 35 et al., Br. J. Cancer, 76(7):852-4 (1997); Friedrich and Weiss, J. Theor. Biol., 187(4):529-40 (1997) and Pieulle, et al., J. Bacteriol., 179(18):5684-92 (1997). Given the physiological importance of redox reactions *in vivo*, efforts are currently being undertaken to identify new, native proteins which are involved in redox

reactions. We describe herein the identification of novel polypeptides which have homology to reductase, designated herein as PRO233 polypeptides.

26. PRO223

5 The carboxypeptidase family of exopeptidases constitutes a diverse group of enzymes that hydrolyze carboxyl-terminal amide bonds in polypeptides, wherein a large number of mammalian tissues produce these enzymes. Many of the carboxypeptidase enzymes that have been identified to date exhibit rather strong cleavage specificities for certain amino acids in polypeptides. For example, carboxypeptidase enzymes have been identified which prefer lysine, arginine, serine or amino acids with either aromatic or branched aliphatic side chains as substrates at the carboxyl terminus of the polypeptide.

10 With regard to the serine carboxypeptidases, such amino acid specific enzymes have been identified from a variety of different mammalian and non-mammalian organisms. The mammalian serine carboxypeptidase enzymes play important roles in many different biological processes including, for example, protein digestion, activation, inactivation, or modulation of peptide hormone activity, and alteration of the physical properties of proteins and enzymes.

15 In light of the physiological importance of the serine carboxypeptidases, efforts are being undertaken by both industry and academia to identify new, native secreted and membrane-bound receptor proteins and specifically novel carboxypeptidases. Many of these efforts are focused on the screening of mammalian recombinant DNA libraries to identify the coding sequences for novel secreted and membrane-bound receptor proteins. We describe herein novel polypeptides having homology to one or more serine carboxypeptidase

20 polypeptides, designated herein as PRO223 polypeptides.

27. PRO235

Plexin was first identified in *Xenopus* tadpole nervous system as a membrane glycoprotein which was shown to mediate cell adhesion via a homophilic binding mechanism in the presence of calcium ions. Strong 25 evolutionary conservation between *Xenopus*, mouse and human homologs of plexin has been observed. [Kaneyama et al., Biochem. And Biophys. Res. Comm. 226: 524-529 (1996)]. Given the physiological importance of cell adhesion mechanisms *in vivo*, efforts are currently being undertaken to identify new, native proteins which are involved in cell adhesion. We describe herein the identification of a novel polypeptide which has homology to plexin, designated herein as PRO235.

30

28. PRO236 and PRO262

β -galactosidase is a well known enzymatic protein which functions to hydrolyze β -galactoside molecules. β -galactosidase has been employed for a variety of different applications, both *in vitro* and *in vivo* and has proven to be an extremely useful research tool. As such, there is an interest in obtaining novel polypeptides which 35 exhibit homology to the β -galactosidase polypeptide.

Given the strong interest in obtaining novel polypeptides having homology to β -galactosidase, efforts are currently being undertaken by both industry and academia to identify new, native β -galactosidase homolog

proteins. Many of these efforts are focused on the screening of mammalian recombinant DNA libraries to identify the coding sequences for novel β -galactosidase-like proteins. Examples of screening methods and techniques are described in the literature [see, for example, Klein et al., Proc. Natl. Acad. Sci., 93:7108-7113 (1996); U.S. Patent No. 5,536,637]. We herein describe novel polypeptides having significant homology to the β -galactosidase enzyme, designated herein as PRO236 and PRO262 polypeptides.

5

29. PRO239

Densin is a glycoprotein which has been isolated from the brain which has all the hallmarks of an adhesion molecule. It is highly concentrated at synaptic sites in the brain and is expressed prominently in dendritic processes in developing neurons. Densin has been characterized as a member of the O-linked sialoglycoproteins. Densin has relevance to medically important processes such as regeneration. Given the physiological importance of synaptic processes and cell adhesion mechanisms *in vivo*, efforts are currently being undertaken to identify new, native proteins which are involved in synaptic machinery and cell adhesion. We describe herein the identification of novel polypeptides which have homology to densin, designated herein as PRO239 polypeptides.

15

30. PRO257

Ebnerin is a cell surface protein associated with von Ebner glands in mammals. Efforts are being undertaken by both industry and academia to identify new, native cell surface receptor proteins and specifically those which possess sequence homology to cell surface proteins such as ebnerin. Many of these efforts are focused on the screening of mammalian recombinant DNA libraries to identify the coding sequences for novel receptor proteins. We herein describe the identification of novel polypeptides having significant homology to the von Ebner's gland-associated protein ebnerin, designated herein as PRO257 polypeptides.

31. PRO260

25 Fucosidases are enzymes that remove fucose residues from fucose containing proteoglycans. In some pathological conditions, such as cancer, rheumatoid arthritis, and diabetes, there is an abnormal fucosylation of serum proteins. Therefore, fucosidases, and proteins having homology to fucosidase, are of importance to the study and abrogation of these conditions. In particular, proteins having homology to the alpha-1-fucosidase precursor are of interest. Fucosidases and fucosidase inhibitors are further described in U.S. Patent Nos. 30 5,637,490, 5,382,709, 5,240,707, 5,153,325, 5,100,797, 5,096,909 and 5,017,704. Studies are also reported in Valk, et al., J. Virol., 71(9):6796 (1997), Aktogu, et al., Monaldi. Arch. Chest Dis. (Italy), 52(2):118 (1997) and Focarelli, et al., Biochem. Biophys. Res. Commun. (U.S.), 234(1):54 (1997).

Efforts are being undertaken by both industry and academia to identify new, native secreted and membrane-bound receptor proteins. Of particular interest are proteins having homology to the alpha-1-fucosidase precursor. Many efforts are focused on the screening of mammalian recombinant DNA libraries to identify the coding sequences for novel secreted and membrane-bound receptor proteins. Examples of screening methods and techniques are described in the literature [see, for example, Klein et al., Proc. Natl. Acad. Sci., 93:7108-

7113 (1996); U.S. Patent No. 5,536,637)].

We herein describe the identification and characterization of novel polypeptides having homology to fucosidases, designated herein as PRO260 polypeptides.

32. **PRO263**

5 CD44 is a cell surface adhesion molecule involved in cell-cell and cell-matrix interactions. Hyaluronic acid, a component of the extracellular matrix is a major ligand. Other ligands include collagen, fibronectin, laminin, chondroitin sulfate, mucosal addressin, serglycin and osteoponin. CD44 is also important in regulating cell traffic, lymph node homing, transmission of growth signals, and presentation of chemokines and growth factors to traveling cells. CD44 surface proteins are associated with metastatic tumors and CD44 has been used
10 as a marker for HIV infection. Certain splice variants are associated with metastasis and poor prognosis of cancer patients. Therefore, molecules having homology with CD44 are of particular interest, as their homology indicates that they may have functions related to those functions of CD44. CD44 is further described in U.S. Patent Nos. 5,506,119, 5,504,194 and 5,108,904; Gerberick, et al., Toxicol. Appl. Pharmacol., 146(1):1 (1997); Wittig, et al., Immunol. Letters (Netherlands), 57(1-3):217 (1997); and Oliveira and Odell, Oral Oncol.
15 (England), 33(4):260 (1997).

Efforts are being undertaken by both industry and academia to identify new, native secreted and membrane-bound receptor proteins, particularly transmembrane proteins with homology to CD44 antigen. Many efforts are focused on the screening of mammalian recombinant DNA libraries to identify the coding sequences for novel secreted and membrane-bound receptor proteins. Examples of screening methods and techniques are
20 described in the literature [see, for example, Klein et al., Proc. Natl. Acad. Sci., 93:7108-7113 (1996); U.S. Patent No. 5,536,637].

We herein describe the identification and characterization of novel polypeptides having homology to CD44 antigen, designated herein as PRO263 polypeptides.

25 33. **PRO270**

Thioredoxins effect reduction-oxidation (redox) state. Many diseases are potentially related to redox state and reactive oxygen species may play a role in many important biological processes. The transcription factors, NF-kappa B and AP-1, are regulated by redox state and are known to affect the expression of a large variety of genes thought to be involved in the pathogenesis of AIDS, cancer, atherosclerosis and diabetic
30 complications. Such proteins may also play a role in cellular antioxidant defense, and in pathological conditions involving oxidative stress such as stroke and inflammation in addition to having a role in apoptosis. Therefore, thioredoxins, and proteins having homology thereto, are of interest to the scientific and medical communities.

We herein describe the identification and characterization of novel polypeptides having homology to thioredoxin, designated herein as PRO270 polypeptides.

35

34. **PRO271**

The proteoglycan link protein is a protein which is intimately associated with various extracellular

matrix proteins and more specifically with proteins such as collagen. For example, one primary component of collagen is a large proteoglycan called aggrecan. This molecule is retained by binding to the glycosaminoglycan hyaluronan through the amino terminal G1 globular domain of the core protein. This binding is stabilized by the proteoglycan link protein which is a protein that is also associated with other tissues containing hyaluronan binding proteoglycans such as versican.

5 Link protein has been identified as a potential target for autoimmune antibodies in individuals who suffer from juvenile rheumatoid arthritis (see Guerassimov et al., *J. Rheumatology* 24(5):959-964 (1997)). As such, there is strong interest in identifying novel proteins having homology to link protein. We herein describe the identification and characterization of novel polypeptides having such homology, designated herein as PRO271 polypeptides.

10

35. PRO272

Reticulocalbin is an endoplasmic reticular protein which may be involved in protein transport and luminal protein processing. Reticulocalbin resides in the lumen of the endoplasmic reticulum, is known to bind calcium, and may be involved in a luminal retention mechanism of the endoplasmic reticulum. It contains six 15 domains of the EF-hand motif associated with high affinity calcium binding. We describe herein the identification and characterization of a novel polypeptide which has homology to the reticulocalbin protein, designated herein as PRO272.

36. PRO294

20 Collagen, a naturally occurring protein, finds wide application in industry. Chemically hydrolyzed natural collagen can be denatured and renatured by heating and cooling to produce gelatin, which is used in photographic and medical, among other applications. Collagen has important properties such as the ability to form interchain aggregates having a conformation designated as a triple helix. We herein describe the identification and characterization of a novel polypeptide which has homology to portions of the collagen 25 molecule, designated herein as PRO294.

37. PRO295

The integrins comprise a supergene family of cell-surface glycoprotein receptors that promote cellular adhesion. Each cell has numerous receptors that define its cell adhesive capabilities. Integrins are involved in 30 a wide variety of interaction between cells and other cells or matrix components. The integrins are of particular importance in regulating movement and function of immune system cells. The platelet IIb/IIIa integrin complex is of particular importance in regulating platelet aggregation. A member of the integrin family, integrin β -6, is expressed on epithelial cells and modulates epithelial inflammation. Another integrin, leucocyte-associated antigen-1 (LFA-1) is important in the adhesion of lymphocytes during an immune response. The integrins are 35 expressed as heterodimers of non-covalently associated alpha and beta subunits. Given the physiological importance of cell adhesion mechanisms *in vivo*, efforts are currently being undertaken to identify new, native proteins which are involved in cell adhesion. We describe herein the identification and characterization of a

novel polypeptide which has homology to integrin, designated herein as PRO295.

38. PRO293

Protein-protein interactions include receptor and antigen complexes and signaling mechanisms. As more is known about the structural and functional mechanisms underlying protein-protein interactions, protein-protein interactions can be more easily manipulated to regulate the particular result of the protein-protein interaction. Thus, the underlying mechanisms of protein-protein interactions are of interest to the scientific and medical community.

All proteins containing leucine-rich repeats are thought to be involved in protein-protein interactions. Leucine-rich repeats are short sequence motifs present in a number of proteins with diverse functions and cellular locations. The crystal structure of ribonuclease inhibitor protein has revealed that leucine-rich repeats correspond to beta-alpha structural units. These units are arranged so that they form a parallel beta-sheet with one surface exposed to solvent, so that the protein acquires an unusual, nonglobular shape. These two features have been indicated as responsible for the protein-binding functions of proteins containing leucine-rich repeats. See, Kobe and Deisenhofer, Trends Biochem. Sci., 19(10):415-421 (Oct. 1994).

15 A study has been reported on leucine-rich proteoglycans which serve as tissue organizers, orienting and ordering collagen fibrils during ontogeny and are involved in pathological processes such as wound healing, tissue repair, and tumor stroma formation. Iozzo, R. V., Crit. Rev. Biochem. Mol. Biol., 32(2):141-174 (1997). Others studies implicating leucine rich proteins in wound healing and tissue repair are De La Salle, C., et al., Vouv. Rev. Fr. Hematol. (Germany), 37(4):215-222 (1995), reporting mutations in the leucine rich motif 20 in a complex associated with the bleeding disorder Bernard-Soulier syndrome and Chlemetson, K. J., Thromb. Haemost. (Germany), 74(1):111-116 (July 1995), reporting that platelets have leucine rich repeats. Another protein of particular interest which has been reported to have leucine-rich repeats is the SLIT protein which has been reported to be useful in treating neuro-degenerative diseases such as Alzheimer's disease, nerve damage such as in Parkinson's disease, and for diagnosis of cancer, see, Artavanistsakonas, S. and Rothberg, J. M., 25 WO9210518-A1 by Yale University. Other studies reporting on the biological functions of proteins having leucine-rich repeats include: Tayar, N., et al., Mol. Cell Endocrinol., (Ireland), 125(1-2):65-70 (Dec. 1996) (gonadotropin receptor involvement); Miura, Y., et al., Nippon Rinsho (Japan), 54(7):1784-1789 (July 1996) (apoptosis involvement); Harris, P. C., et al., J. Am. Soc. Nephrol., 6(4):1125-1133 (Oct. 1995) (kidney disease involvement); and Ruoslahti, E. I., et al., WO9110727-A by La Jolla Cancer Research Foundation 30 (decorin binding to transforming growth factor β involvement for treatment for cancer, wound healing and scarring).

Efforts are therefore being undertaken by both industry and academia to identify new proteins having leucine rich repeats to better understand protein-protein interactions. Of particular interest are those proteins having leucine rich repeats and homology to known neuronal leucine rich repeat proteins. Many efforts are 35 focused on the screening of mammalian recombinant DNA libraries to identify the coding sequences for novel secreted and membrane-bound proteins having leucine rich repeats. Examples of screening methods and techniques are described in the literature [see, for example, Klein et al., Proc. Natl. Acad. Sci., 93:7108-7113

(1996); U.S. Patent No. 5,536,637)].

We describe herein the identification and characterization of a novel polypeptide which has homology to leucine rich repeat proteins, designated herein as PRO293.

39. PRO247

5 Protein-protein interactions include receptor and antigen complexes and signaling mechanisms. As more is known about the structural and functional mechanisms underlying protein-protein interactions, protein-protein interactions can be more easily manipulated to regulate the particular result of the protein-protein interaction. Thus, the underlying mechanisms of protein-protein interactions are of interest to the scientific and medical community.

10 All proteins containing leucine-rich repeats are thought to be involved in protein-protein interactions. Leucine-rich repeats are short sequence motifs present in a number of proteins with diverse functions and cellular locations. The crystal structure of ribonuclease inhibitor protein has revealed that leucine-rich repeats correspond to beta-alpha structural units. These units are arranged so that they form a parallel beta-sheet with one surface exposed to solvent, so that the protein acquires an unusual, nonglobular shape. These two features
15 have been indicated as responsible for the protein-binding functions of proteins containing leucine-rich repeats. See, Kobe and Deisenhofer, Trends Biochem. Sci., 19(10):415-421 (Oct. 1994).

A study has been reported on leucine-rich proteoglycans which serve as tissue organizers, orienting and ordering collagen fibrils during ontogeny and are involved in pathological processes such as wound healing, tissue repair, and tumor stroma formation. Iozzo, R. V., Crit. Rev. Biochem. Mol. Biol., 32(2):141-174
20 (1997). Others studies implicating leucine rich proteins in wound healing and tissue repair are De La Salle, C., et al., Vouv. Rev. Fr. Hematol. (Germany), 37(4):215-222 (1995), reporting mutations in the leucine rich motif in a complex associated with the bleeding disorder Bernard-Soulier syndrome and Chlernetson, K. J., Thromb. Haemost. (Germany), 74(1):111-116 (July 1995), reporting that platelets have leucine rich repeats. Another protein of particular interest which has been reported to have leucine-rich repeats is the SLIT protein which has
25 been reported to be useful in treating neuro-degenerative diseases such as Alzheimer's disease, nerve damage such as in Parkinson's disease, and for diagnosis of cancer, see, Artavanistsakonas, S. and Rothberg, J. M., WO9210518-A1 by Yale University. Other studies reporting on the biological functions of proteins having leucine-rich repeats include: Tayar, N., et al., Mol. Cell Endocrinol., (Ireland), 125(1-2):65-70 (Dec. 1996)
30 (gonadotropin receptor involvement); Miura, Y., et al., Nippon Rinsho (Japan), 54(7):1784-1789 (July 1996) (apoptosis involvement); Harris, P. C., et al., J. Am. Soc. Nephrol., 6(4):1125-1133 (Oct. 1995) (kidney disease involvement); and Ruoslahti, E. I., et al., WO9110727-A by La Jolla Cancer Research Foundation (decorin binding to transforming growth factor β involvement for treatment for cancer, wound healing and scarring).

35 Densin is a glycoprotein which has been isolated from the brain which has all the hallmarks of an adhesion molecule. It is highly concentrated at synaptic sites in the brain and is expressed prominently in dendritic processes in developing neurons. Densin has been characterized as a member of the O-linked sialoglycoproteins. Densin has relevance to medically important processes such as regeneration. Given the

physiological importance of synaptic processes and cell adhesion mechanisms *in vivo*, efforts are currently being undertaken to identify new, native proteins which are involved in synaptic machinery and cell adhesion. Densin is further described in Kennedy, M.B., Trends Neurosci. (England), 20(6):264 (1997) and Apperson, et al., J. Neurosci., 16(21):6839 (1996).

Efforts are therefore being undertaken by both industry and academia to identify new proteins having 5 leucine rich repeats to better understand protein-protein interactions. Of particular interest are those proteins having leucine rich repeats and homology to known proteins having leucine rich repeats such as KIAA0231 and densin. Many efforts are focused on the screening of mammalian recombinant DNA libraries to identify the coding sequences for novel secreted and membrane-bound proteins having leucine rich repeats. Examples of screening methods and techniques are described in the literature [see, for example, Klein et al., Proc. Natl. 10 Acad. Sci., 93:7108-7113 (1996); U.S. Patent No. 5,536,637)].

We describe herein the identification and characterization of a novel polypeptide which has homology to leucine rich repeat proteins, designated herein as PRO247.

40. PRO302, PRO303, PRO304, PRO307 and PRO343

15 Proteases are enzymatic proteins which are involved in a large number of very important biological processes in mammalian and non-mammalian organisms. Numerous different protease enzymes from a variety of different mammalian and non-mammalian organisms have been both identified and characterized. The mammalian protease enzymes play important roles in many different biological processes including, for example, protein digestion, activation, inactivation, or modulation of peptide hormone activity, and alteration of the 20 physical properties of proteins and enzymes.

In light of the important physiological roles played by protease enzymes, efforts are currently being undertaken by both industry and academia to identify new, native protease homologs. Many of these efforts are focused on the screening of mammalian recombinant DNA libraries to identify the coding sequences for novel secreted and membrane-bound receptor proteins. Examples of screening methods and techniques are described 25 in the literature [see, for example, Klein et al., Proc. Natl. Acad. Sci., 93:7108-7113 (1996); U.S. Patent No. 5,536,637)]. We herein describe the identification of novel polypeptides having homology to various protease enzymes, designated herein as PRO302, PRO303, PRO304, PRO307 and PRO343 polypeptides.

41. PRO328

30 The GLIP protein family has been characterized as comprising zinc-finger proteins which play important roles in embryogenesis. These proteins may function as transcriptional regulatory proteins and are known to be amplified in a subset of human tumors. Glioma pathogenesis protein is structurally related to a group of plant pathogenesis-related proteins. It is highly expressed in glioblastoma. See US Pat. Nos. 5,582,981 (issued Dec. 10, 1996) and 5,322,801 (issued June 21, 1996), Ellington, A.D. et al., Nature, 346:818 (1990), Grindley, J.C. 35 et al., Dev. Biol., 188(2):337 (1997), Marine, J.C. et al., Mech. Dev., 63(2):211 (1997). The CRISP or cysteine rich secretory protein family are a group of proteins which are also structurally related to a group of plant pathogenesis proteins. [Schwidetzky, U., Biochem. J., 321:325 (1997), Pfisterer, P., Mol. Cell Biol.,

16(11):6160 (1996), Kratzschmar, J., Eur. J. Biochem., 236(3):827 (1996)]. We describe herein the identification of a novel polypeptide which has homology to GLIP and CRISP, designated herein as PRO328 polypeptides.

42. PRO335, PRO331 and PRO326

5 Protein-protein interactions include receptor and antigen complexes and signaling mechanisms. As more is known about the structural and functional mechanisms underlying protein-protein interactions, protein-protein interactions can be more easily manipulated to regulate the particular result of the protein-protein interaction. Thus, the underlying mechanisms of protein-protein interactions are of interest to the scientific and medical community.

10 All proteins containing leucine-rich repeats are thought to be involved in protein-protein interactions. Leucine-rich repeats are short sequence motifs present in a number of proteins with diverse functions and cellular locations. The crystal structure of ribonuclease inhibitor protein has revealed that leucine-rich repeats correspond to beta-alpha structural units. These units are arranged so that they form a parallel beta-sheet with one surface exposed to solvent, so that the protein acquires an unusual, nonglobular shape. These two features 15 have been indicated as responsible for the protein-binding functions of proteins containing leucine-rich repeats. See, Kobe and Deisenhofer, Trends Biochem. Sci., 19(10):415-421 (Oct. 1994).

A study has been reported on leucine-rich proteoglycans which serve as tissue organizers, orienting and ordering collagen fibrils during ontogeny and are involved in pathological processes such as wound healing, tissue repair, and tumor stroma formation. Iozzo, R. V., Crit. Rev. Biochem. Mol. Biol., 32(2):141-174 20 (1997). Others studies implicating leucine rich proteins in wound healing and tissue repair are De La Salle, C., et al., Vouv. Rev. Fr. Hematol. (Germany), 37(4):215-222 (1995), reporting mutations in the leucine rich motif in a complex associated with the bleeding disorder Bernard-Soulier syndrome, Chlemetson, K. J., Thromb. Haemost. (Germany), 74(1):111-116 (July 1995), reporting that platelets have leucine rich repeats and Ruoslahti, E. I., et al., WO9110727-A by La Jolla Cancer Research Foundation reporting that decorin binding to 25 transforming growth factor β has involvement in a treatment for cancer, wound healing and scarring. Related by function to this group of proteins is the insulin like growth factor (IGF), in that it is useful in wound-healing and associated therapies concerned with re-growth of tissue, such as connective tissue, skin and bone; in promoting body growth in humans and animals; and in stimulating other growth-related processes. The acid labile subunit of IGF (ALS) is also of interest in that it increases the half-life of IGF and is part of the IGF complex in vivo.

30 Another protein which has been reported to have leucine-rich repeats is the SLIT protein which has been reported to be useful in treating neuro-degenerative diseases such as Alzheimer's disease, nerve damage such as in Parkinson's disease, and for diagnosis of cancer, see, Artavanistsakonas, S. and Rothberg, J. M., WO9210518-A1 by Yale University. Of particular interest is LIG-1, a membrane glycoprotein that is expressed specifically in glial cells in the mouse brain, and has leucine rich repeats and immunoglobulin-like domains. 35 Suzuki, et al., J. Biol. Chem. (U.S.), 271(37):22522 (1996). Other studies reporting on the biological functions of proteins having leucine rich repeats include: Tayar, N., et al., Mol. Cell Endocrinol., (Ireland), 125(1-2):65-70 (Dec. 1996) (gonadotropin receptor involvement); Miura, Y., et al., Nippon Rinsho (Japan), 54(7):1784-1789

(July 1996) (apoptosis involvement); Harris, P. C., et al., J. Am. Soc. Nephrol., 6(4):1125-1133 (Oct. 1995) (kidney disease involvement).

Efforts are therefore being undertaken by both industry and academia to identify new proteins having leucine rich repeats to better understand protein-protein interactions. Of particular interest are those proteins having leucine rich repeats and homology to known proteins having leucine rich repeats such as LIG-1, ALS and decorin. Many efforts are focused on the screening of mammalian recombinant DNA libraries to identify the coding sequences for novel secreted and membrane-bound proteins having leucine rich repeats. Examples of screening methods and techniques are described in the literature [see, for example, Klein et al., Proc. Natl. Acad. Sci., 93:7108-7113 (1996); U.S. Patent No. 5,536,637].

We describe herein the identification and characterization of novel polypeptides which have homology to proteins of the leucine rich repeat superfamily, designated herein as PRO335, PRO331 and PRO326 polypeptides.

43. PRO332

Secreted proteins comprising a repeat characterized by an arrangement of conserved leucine residues (leucine-rich repeat motif) have diverse biological roles. Certain proteoglycans, such as biglycan, fibromodulin and decorin, are, for example, characterized by the presence of a leucine-rich repeat of about 24 amino acids [Ruoslahti, Ann. Rev. Cell. Biol., 4 229-255 (1988); Oldberg *et al.*, EMBO J., 8, 2601-2604 (1989)]. In general, proteoglycans are believed to play a role in regulating extracellular matrix, cartilage or bone function. The proteoglycan decorin binds to collagen type I and II and affects the rate of fibril formation. Fibromodulin also binds collagen and delays fibril formation. Both fibromodulin and decorin inhibit the activity of transforming growth factor beta (TGF- β) (U.S. Patent No. 5,583,103 issued December 10, 1996). TGF- β is known to play a key role in the induction of extracellular matrix and has been implicated in the development of fibrotic diseases, such as cancer and glomerulonephritis. Accordingly, proteoglycans have been proposed for the treatment of fibrotic cancer, based upon their ability to inhibit TGF- β 's growth stimulating activity on the cancer cell. Proteoglycans have also been described as potentially useful in the treatment of other proliferative pathologies, including rheumatoid arthritis, arteriosclerosis, adult respiratory distress syndrome, cirrhosis of the liver, fibrosis of the lungs, post-myocardial infarction, cardiac fibrosis, post-angioplasty restenosis, renal interstitial fibrosis and certain dermal fibrotic conditions, such as keloids and scarring, which might result from burn injuries, other invasive skin injuries, or cosmetic or reconstructive surgery (U.S. Patent No. 5,654,270, issued August 5, 1997).

We describe herein the identification and characterization of novel polypeptides which have homology to proteins of the leucine rich repeat superfamily, designated herein as PRO332 polypeptides.

44. PRO334

Microfibril bundles and proteins found in association with these bundles, particularly attachment molecules, are of interest in the field of dermatology, particularly in the study of skin which has been damaged from aging, injuries or the sun. Fibrillin microfibrils define the continuous elastic network of skin, and are

present in dermis as microfibril bundles devoid of measurable elastin extending from the dermal-epithelial junction and as components of the thick elastic fibres present in the deep reticular dermis. Moreover, Marfan syndrome has been linked to mutations which interfere with multimerization of fibrillin monomers or other connective tissue elements.

5 Fibulin-1 is a modular glycoprotein with amino-terminal anaphlatoxin-like modules followed by nine epidermal growth factor (EGF)-like modules and, depending on alternative splicing, four possible carboxyl termini. Fibulin-2 is a novel extracellular matrix protein frequently found in close association with microfibrils containing either fibronectin or fibrillin. Thus, fibrillin, fibulin, and molecules related thereto are of interest, particularly for the use of preventing skin from being damaged from aging, injuries or the sun, or for restoring skin damaged from same. Moreover, these molecules are generally of interest in the study of connective tissue
10 and attachment molecules and related mechanisms. Fibrillin, fibulin and related molecules are further described in Adams, et al., J. Mol. Biol., 272(2):226-36 (1997); Kiely and Shuttleworth, Microsc. Res. Tech., 38(4):413-27 (1997); and Child, J. Card. Surg., 12(2Supp.):131-5 (1997).

15 Currently, efforts are being undertaken by both industry and academia to identify new, native secreted and membrane-bound receptor proteins, particularly secreted proteins which have homology to fibulin and fibrillin. Many efforts are focused on the screening of mammalian recombinant DNA libraries to identify the coding sequences for novel secreted and membrane-bound receptor proteins. Examples of screening methods and techniques are described in the literature [see, for example, Klein et al., Proc. Natl. Acad. Sci., 93:7108-7113 (1996); U.S. Patent No. 5,536,637)].

20 We herein describe the identification and characterization of novel polypeptides having homology to fibulin and fibrillin, designated herein as PRO334 polypeptides.

45. PRO346

25 The widespread occurrence of cancer has prompted the devotion of considerable resources and discovering new treatments of treatment. One particular method involves the creation of tumor or cancer specific monoclonal antibodies (mAbs) which are specific to tumor antigens. Such mAbs, which can distinguish between normal and cancerous cells are useful in the diagnosis, prognosis and treatment of the disease. Particular antigens are known to be associated with neoplastic diseases, such as colorectal and breast cancer. Since colon cancer is a widespread disease, early diagnosis and treatment is an important medical goal. Diagnosis and treatment of cancer can be implemented using monoclonal antibodies (mAbs) specific therefore having
30 fluorescent, nuclear magnetic or radioactive tags. Radioactive genes, toxins and/or drug tagged mAbs can be used for treatment *in situ* with minimal patient description.

35 Carcinoembryonic antigen (CEA) is a glycoprotein found in human colon cancer and the digestive organs of a 2-6 month human embryos. CEA is a known human tumor marker and is widely used in the diagnosis of neoplastic diseases, such as colon cancer. For example, when the serum levels of CEA are elevated in a patient, a drop of CEA levels after surgery would indicate the tumor resection was successful. On the other hand, a subsequent rise in serum CEA levels after surgery would indicate that metastases of the original tumor may have formed or that new primary tumors may have appeared. CEA may also be a target for mAb, antisense

nucleotides

46. **PRO268**

Protein disulfide isomerase is an enzymatic protein which is involved in the promotion of correct refolding of proteins through the establishment of correct disulfide bond formation. Protein disulfide isomerase 5 was initially identified based upon its ability to catalyze the renaturation of reduced denatured RNase (Goldberger et al., *J. Biol. Chem.* 239:1406-1410 (1964) and Epstein et al., *Cold Spring Harbor Symp. Quant. Biol.* 28:439-449 (1963)). Protein disulfide isomerase has been shown to be a resident enzyme of the endoplasmic reticulum which is retained in the endoplasmic reticulum via a -KDEL or -HDEL amino acid sequence at its C-terminus.

10 Given the importance of disulfide bond-forming enzymes and their potential uses in a number of different applications, for example in increasing the yield of correct refolding of recombinantly produced proteins, efforts are currently being undertaken by both industry and academia to identify new, native proteins having homology to protein disulfide isomerase. Many of these efforts are focused on the screening of mammalian recombinant DNA libraries to identify the coding sequences for novel protein disulfide isomerase 15 homologs. We herein describe a novel polypeptide having homology to protein disulfide isomerase, designated herein as PRO268.

47. **PRO330**

Prolyl 4-hydroxylase is an enzyme which functions to post-translationally hydroxylate proline residues 20 at the Y position of the amino acid sequence Gly-X-Y, which is a repeating three amino acid sequence found in both collagen and procollagen. Hydroxylation of proline residues at the Y position of the Gly-X-Y amino acid triplet to form 4-hydroxyproline residues at those positions is required before newly synthesized collagen polypeptide chains may fold into their proper three-dimensional triple-helical conformation. If hydroxylation does not occur, synthesized collagen polypeptides remain non-helical, are poorly secreted by cells and cannot 25 assemble into stable functional collagen fibrils. Vuorio et al., *Proc. Natl. Acad. Sci. USA* 89:7467-7470 (1992). Prolyl 4-hydroxylase is comprised of at least two different polypeptide subunits, alpha and beta.

Efforts are being undertaken by both industry and academia to identify new, native secreted and membrane-bound receptor proteins. Many efforts are focused on the screening of mammalian recombinant DNA 30 libraries to identify the coding sequences for novel secreted and membrane-bound receptor proteins. Examples of screening methods and techniques are described in the literature [see, for example, Klein et al., *Proc. Natl. Acad. Sci.*, 93:7108-7113 (1996); U.S. Patent No. 5,536,637]. Based upon these efforts, Applicants have herein identified and describe a novel polypeptide having homology to the alpha subunit of prolyl 4-hydroxylase, designated herein as PRO330.

35 48. **PRO339 and PRO310**

Fringe is a protein which specifically blocks serrate-mediated activation of notch in the dorsal compartment of the *Drosophila* wing imaginal disc. Fleming, et al., *Development*, 124(15):2973-81 (1997).

Therefore, fringe is of interest for both its role in development as well as its ability to regulate serrate, particularly serrate's signaling abilities. Also of interest are novel polypeptides which may have a role in development and/or the regulation of serrate-like molecules. Of particular interest are novel polypeptides having homology to fringe as identified and described herein, designated herein as PRO339 and PRO310 polypeptides.

5 49. PRO244

Lectins are a class of proteins comprising a region that binds carbohydrates specifically and non-covalently. Numerous lectins have been identified in higher animals, both membrane-bound and soluble, and have been implicated in a variety of cell-recognition phenomena and tumor metastasis.

Most lectins can be classified as either C-type (calcium-dependent) or S-type (thiol-dependent).

10 Lectins are thought to play a role in regulating cellular events that are initiated at the level of the plasma membrane. For example, plasma membrane associated molecules are involved in the activation of various subsets of lymphoid cells, e.g. T-lymphocytes, and it is known that cell surface molecules are responsible for activation of these cells and consequently their response during an immune reaction.

A particular group of cell adhesion molecules, selectins, belong in the superfamily of C-type lectins.

15 This group includes L-selectin (peripheral lymph node homing receptor (pNHR), LEC-CAM-1, LAM-1, gp90^{MEL}, gp100^{MEL}, gp110^{MEL}, MEL-14 antigen, Leu-8 antigen, TQ-1 antigen, DREG antigen), E-selectin (LEC-CAM-2, LECAM-2, ELAM-1), and P-selectin (LEC-CAM-3, LECAM-3, GMP-140, PADGEM). The structure of selectins consists of a C-type lectin (carbohydrate binding) domain, an epidermal growth factor-like (EGF-like) motif, and variable numbers of complement regulatory (CR) motifs. Selectins are associated with 20 leukocyte adhesion, e.g. the attachment of neutrophils to venular endothelial cells adjacent to inflammation (E-selectin), or with the trafficking of lymphocytes from blood to secondary lymphoid organs, e.g. lymph nodes and Peyer's patches (L-selectin).

Another exemplary lectin is the cell-associated macrophage antigen, Mac-2 that is believed to be involved in cell adhesion and immune responses. Macrophages also express a lectin that recognizes Tn Ag, a 25 human carcinoma-associated epitope.

Another C-type lectin is CD95 (Fas antigen/APO-1) that is an important mediator of immunologically relevant regulated or programmed cell death (apoptosis). "Apoptosis" is a non-necrotic cell death that takes place in metazoan animal cells following activation of an intrinsic cell suicide program. The cloning of Fas antigen is described in PCT publication WO 91/10448, and European patent application EP510691. The mature 30 Fas molecule consists of 319 amino acids of which 157 are extracellular, 17 constitute the transmembrane domain, and 145 are intracellular. Increased levels of Fas expression at T cell surface have been associated with tumor cells and HIV-infected cells. Ligation of CD95 triggers apoptosis in the presence of interleukin-1 (IL-2).

C-type lectins also include receptors for oxidized low-density lipoprotein (LDL). This suggests a possible role in the pathogenesis of atherosclerosis.

35 We herein describe the identification and characterization of novel polypeptides having homology to C-type lectins, designated herein as PRO244 polypeptides.

SUMMARY OF THE INVENTION**1. PRO211 and PRO217**

Applicants have identified cDNA clones that encode novel polypeptides having homology to EGF, designated in the present application as "PRO211" and "PRO217" polypeptides.

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding 5 a PRO211 or PRO217 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding EGF-like homologue PRO211 and PRO217 polypeptides of Fig. 2 (SEQ ID NO:2) and/or 4 (SEQ ID NO:4) indicated in Fig. 1 (SEQ ID NO:1) and/or Fig. 3 (SEQ ID NO:3), respectively, or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

10 In another embodiment, the invention provides isolated PRO211 and PRO217 EGF-like homologue PRO211 and PRO217 polypeptides. In particular, the invention provides isolated native sequence PRO211 and PRO217 EGF-like homologue polypeptides, which in one embodiment, includes an amino acid sequence comprising residues: 1 to 353 of Fig. 2 (SEQ ID NO:2) or (2) 1 to 379 of Fig. 4 (SEQ ID NO: 4).

15 2. PRO230

Applicants have identified a cDNA clone that encodes a novel polypeptide, wherein the polypeptide is designated in the present application as "PRO230".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO230 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO230 20 polypeptide having amino acid residues 1 through 467 of Figure 6 (SEQ ID NO:12), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO230 polypeptide. In particular, the invention provides isolated native sequence PRO230 polypeptide, which in one embodiment, includes an amino 25 acid sequence comprising residues 1 through 467 of Figure 6 (SEQ ID NO:12).

In another embodiment, the invention provides an expressed sequence tag (EST) comprising the nucleotide sequence of SEQ ID NO:13 (Figure 7) which is herein designated as DNA20088.

30 3. PRO232

Applicants have identified a cDNA clone that encodes a novel polypeptide, wherein the polypeptide is designated in the present application as "PRO232".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO232 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO232 35 polypeptide having amino acid residues 1 to 114 of Figure 9 (SEQ ID NO:18), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO232 polypeptide. In particular, the invention provides isolated native sequence PRO232 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 114 of Figure 9 (SEQ ID NO:18).

4. **PRO187**

5 Applicants have identified a cDNA clone that encodes a novel polypeptide, designated in the present application as "PRO187".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO187 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO187 polypeptide of Figure 11 (SEQ ID NO:23), or is complementary to such encoding nucleic acid sequence, and 10 remains stably bound to it under at least moderate, and optionally, under high stringency conditions. In another aspect, the invention provides a nucleic acid comprising the coding sequence of Figure 10 (SEQ ID NO:22) or its complement. In another aspect, the invention provides a nucleic acid of the full length protein of clone DNA27864-1155, deposited with the ATCC under accession number ATCC 209375, alternatively the coding sequence of clone DNA27864-1155, deposited under accession number ATCC 209375.

15 In yet another embodiment, the invention provides isolated PRO187 polypeptide. In particular, the invention provides isolated native sequence PRO187 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 205 of Figure 11 (SEQ ID NO:23). Alternatively, the invention provides a polypeptide encoded by the nucleic acid deposited under accession number ATCC 209375.

20 5. **PRO265**

Applicants have identified a cDNA clone that encodes a novel polypeptide, wherein the polypeptide is designated in the present application as "PRO265".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO265 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO265 polypeptide having amino acid residues 1 to 660 of Figure 13 (SEQ ID NO:28), or is complementary to such 25 encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO265 polypeptide. In particular, the invention provides isolated native sequence PRO265 polypeptide, which in one embodiment, includes an amino 30 acid sequence comprising residues 1 to 660 of Figure 13 (SEQ ID NO:28). An additional embodiment of the present invention is directed to an isolated extracellular domain of a PRO265 polypeptide.

6. **PRO219**

Applicants have identified a cDNA clone that encodes a novel polypeptide, wherein the polypeptide is 35 designated in the present application as "PRO219".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO219 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO219

polypeptide having amino acid residues 1 to 915 of Figure 15 (SEQ ID NO:34), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

5 In another embodiment, the invention provides isolated PRO219 polypeptide. In particular, the invention provides isolated native sequence PRO219 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 915 of Figure 15 (SEQ ID NO:34).

7. **PRO246**

Applicants have identified a cDNA clone that encodes a novel polypeptide, wherein the polypeptide is designated in the present application as "PRO246".

10 In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO246 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO246 polypeptide having amino acid residues 1 to 390 of Figure 17 (SEQ ID NO:39), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

15 In another embodiment, the invention provides isolated PRO246 polypeptide. In particular, the invention provides isolated native sequence PRO246 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 390 of Figure 17 (SEQ ID NO:39). An additional embodiment of the present invention is directed to an isolated extracellular domain of a PRO246 polypeptide.

20 8. **PRO228**

Applicants have identified a cDNA clone that encodes a novel polypeptide having homology to CD97, EMR1 and Iatrophilin, wherein the polypeptide is designated in the present application as "PRO228".

25 In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO228 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO228 polypeptide having amino acid residues 1 to 690 of Figure 19 (SEQ ID NO:49), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

30 In another embodiment, the invention provides isolated PRO228 polypeptide. In particular, the invention provides isolated native sequence PRO228 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 690 of Figure 19 (SEQ ID NO:49). An additional embodiment of the present invention is directed to an isolated extracellular domain of a PRO228 polypeptide.

In another embodiment, the invention provides an expressed sequence tag (EST) comprising the nucleotide sequence of SEQ ID NO:50, designated herein as DNA21951.

35 9. **PRO533**

Applicants have identified a cDNA clone (DNA49435-1219) that encodes a novel polypeptide, designated in the present application as PRO533.

In one embodiment, the invention provides an isolated nucleic acid molecule having at least about 80% sequence identity to (a) a DNA molecule encoding a PRO533 polypeptide comprising the sequence of amino acids 23 to 216 of Figure 22 (SEQ ID NO:59), or (b) the complement of the DNA molecule of (a). The sequence identity preferably is about 85%, more preferably about 90%, most preferably about 95%. In one aspect, the isolated nucleic acid has at least about 80%, preferably at least about 85%, more preferably at least 5 about 90%, and most preferably at least about 95% sequence identity with a polypeptide having amino acid residues 23 to 216 of Figure 22 (SEQ ID NO:59). Preferably, the highest degree of sequence identity occurs within the secreted portion (amino acids 23 to 216 of Figure 22, SEQ ID NO:59). In a further embodiment, the isolated nucleic acid molecule comprises DNA encoding a PRO533 polypeptide having amino acid residues 1 to 216 of Figure 22 (SEQ ID NO:59), or is complementary to such encoding nucleic acid sequence, and remains 10 stably bound to it under at least moderate, and optionally, under high stringency conditions. In another aspect, the invention provides a nucleic acid of the full length protein of clone DNA49435-1219, deposited with the ATCC under accession number ATCC 209480.

In yet another embodiment, the invention provides isolated PRO533 polypeptide. In particular, the invention provides isolated native sequence PRO533 polypeptide, which in one embodiment, includes an amino 15 acid sequence comprising residues 23 to 216 of Figure 22 (SEQ ID NO:59). Native PRO533 polypeptides with or without the native signal sequence (amino acids 1 to 22 in Figure 22 (SEQ ID NO:59)), and with or without the initiating methionine are specifically included. Alternatively, the invention provides a PRO533 polypeptide encoded by the nucleic acid deposited under accession number ATCC 209480.

20 10. PRO245

Applicants have identified a cDNA clone that encodes a novel polypeptide, wherein the polypeptide is designated in the present application as "PRO245".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO245 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO245 25 polypeptide having amino acid residues 1 to 312 of Fig. 24 (SEQ ID NO:64), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO245 polypeptide. In particular, the invention provides isolated native sequence PRO245 polypeptide, which in one embodiment, includes an amino 30 acid sequence comprising residues 1 to 312 of Figure 24 (SEQ ID NO:64).

11. PRO220, PRO221 and PRO227

Applicants have identified cDNA clones that each encode novel polypeptides, all having leucine rich repeats. These polypeptides are designated in the present application as PRO220, PRO221 and PRO227.

35 In one embodiment, the invention provides isolated nucleic acid molecules comprising DNA respectively encoding PRO220, PRO221 and PRO227, respectively. In one aspect, provided herein is an isolated nucleic acid comprises DNA encoding the PRO220 polypeptide having amino acid residues 1 through 708 of Figure 26

(SEQ ID NO:69), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions. Also provided herein is an isolated nucleic acid comprises DNA encoding the PRO221 polypeptide having amino acid residues 1 through 259 of Figure 28 (SEQ ID NO:71), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions. Moreover, also provided 5 herein is an isolated nucleic acid comprises DNA encoding the PRO227 polypeptide having amino acid residues 1 through 620 of Figure 30 (SEQ ID NO:73), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO220, PRO221 and PRO227 polypeptides. In particular, provided herein is the isolated native sequence for the PRO220 polypeptide, which in one 10 embodiment, includes an amino acid sequence comprising residues 1 to 708 of Figure 26 (SEQ ID NO:69). Additionally provided herein is the isolated native sequence for the PRO221 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 259 of Figure 28 (SEQ ID NO:71). Moreover, provided herein is the isolated native sequence for the PRO227 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 620 of Figure 30 (SEQ ID NO:73).

15

12. PRO258

Applicants have identified a cDNA clone that encodes a novel polypeptide having homology to CRTAM and poliovirus receptor precursors, wherein the polypeptide is designated in the present application as "PRO258".

20 In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO258 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO258 polypeptide having amino acid residues 1 to 398 of Figure 32 (SEQ ID NO:84), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

25 In another embodiment, the invention provides isolated PRO258 polypeptide. In particular, the invention provides isolated native sequence PRO258 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 398 of Figure 32 (SEQ ID NO:84). An additional embodiment of the present invention is directed to an isolated extracellular domain of a PRO258 polypeptide.

30

13. PRO266

Applicants have identified a cDNA clone that encodes a novel polypeptide, wherein the polypeptide is designated in the present application as "PRO266".

35 In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO266 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO266 polypeptide having amino acid residues 1 to 696 of Figure 34 (SEQ ID NO:91), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO266 polypeptide. In particular, the invention provides isolated native sequence PRO266 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 696 of Figure 34 (SEQ ID NO:91).

14. PRO269

5 Applicants have identified a cDNA clone that encodes a novel polypeptide, wherein the polypeptide is designated in the present application as PRO269.

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO269 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO269 polypeptide having amino acid residues 1 to 490 of Fig. 36 (SEQ ID NO:96), or is complementary to such 10 encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO269 polypeptide. In particular, the invention provides isolated native sequence PRO269 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 490 of Figure 36 (SEQ ID NO:96). An additional embodiment of the 15 present invention is directed to an isolated extracellular domain of a PRO269 polypeptide.

15. PRO287

Applicants have identified a cDNA clone that encodes a novel polypeptide, wherein the polypeptide is designated in the present application as "PRO287".

20 In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO287 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO287 polypeptide having amino acid residues 1 to 415 of Fig. 38 (SEQ ID NO:104), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

25 In another embodiment, the invention provides isolated PRO287 polypeptide. In particular, the invention provides isolated native sequence PRO287 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 415 of Figure 38 (SEQ ID NO:104).

16. PRO214

30 Applicants have identified a cDNA clone that encodes a novel polypeptide, designated in the present application as "PRO214".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO214 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO214 polypeptide of Fig. 40 (SEQ ID NO:109), or is complementary to such encoding nucleic acid sequence, and 35 remains stably bound to it under at least moderate, and optionally, under high stringency conditions. In another aspect, the invention provides a nucleic acid comprising the coding sequence of Fig. 39 (SEQ ID NO:108) or its complement. In another aspect, the invention provides a nucleic acid of the full length protein of clone

DNA32286-1191, deposited with ATCC under accession number ATCC 209385.

In yet another embodiment, the invention provides isolated PRO214 polypeptide. In particular, the invention provides isolated native sequence PRO214 polypeptide, which in one embodiment, includes an amino acid sequence comprising the residues of Figure 40 (SEQ ID NO:109). Alternatively, the invention provides a polypeptide encoded by the nucleic acid deposited under accession number ATCC 209385.

5

17. PRO317

Applicants have identified a cDNA clone that encodes a novel polypeptide, designated in the present application as "PRO317".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding 10 PRO317 polypeptide. In one aspect, the isolated nucleic acid comprises DNA (SEQ ID NO:113) encoding PRO317 polypeptide having amino acid residues 1 to 366 of Fig. 42, or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO317 polypeptide. In particular, the 15 invention provides isolated native-sequence PRO317 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 366 of Figure 42 (SEQ ID NO:114).

In yet another embodiment, the invention supplies a method of detecting the presence of PRO317 in a sample, the method comprising:

20 a) contacting a detectable anti-PRO317 antibody with a sample suspected of containing PRO317; and b) detecting binding of the antibody to the sample; wherein the sample is selected from the group consisting of a body fluid, a tissue sample, a cell extract, and a cell culture medium.

In a still further embodiment a method is provided for determining the presence of PRO317 mRNA in a sample, the method comprising:

25 a) contacting a sample suspected of containing PRO317 mRNA with a detectable nucleic acid probe that hybridizes under moderate to stringent conditions to PRO317 mRNA; and b) detecting hybridization of the probe to the sample.

Preferably, in this method the sample is a tissue sample and the detecting step is by *in situ* hybridization, or the sample is a cell extract and detection is by Northern analysis.

Further, the invention provides a method for treating a PRO317-associated disorder comprising 30 administering to a mammal an effective amount of the PRO317 polypeptide or a composition thereof containing a carrier, or with an effective amount of a PRO317 agonist or PRO317 antagonist, such as an antibody which binds specifically to PRO317.

18. PRO301

35 Applicants have identified a cDNA clone (DNA40628-1216) that encodes a novel polypeptide, designated in the present application as "PRO301".

In one embodiment, the invention provides an isolated nucleic acid molecule having at least about 80% sequence identity to (a) a DNA molecule encoding a PRO301 polypeptide comprising the sequence of amino acids 28 to 258 of Fig. 44 (SEQ ID NO:119), or (b) the complement of the DNA molecule of (a). The sequence identity preferably is about 85%, more preferably about 90%, most preferably about 95%. In one aspect, the isolated nucleic acid has at least about 80%, preferably at least about 85%, more preferably at least about 90%, 5 and most preferably at least about 95% sequence identity with a polypeptide having amino acid residues 28 to 258 of Fig. 44 (SEQ ID NO:119). Preferably, the highest degree of sequence identity occurs within the extracellular domains (amino acids 28 to 258 of Fig. 44, SEQ ID NO:119). In a further embodiment, the isolated nucleic acid molecule comprises DNA encoding a PRO301 polypeptide having amino acid residues 28 to 299 of Fig. 44 (SEQ ID NO:119), or is complementary to such encoding nucleic acid sequence, and remains 10 stably bound to it under at least moderate, and optionally, under high stringency conditions. In another aspect, the invention provides a nucleic acid of the full length protein of clone DNA40628-1216, deposited with the ATCC under accession number ATCC 209432, alternatively the coding sequence of clone DNA40628-1216, deposited under accession number ATCC 209432.

In yet another embodiment, the invention provides isolated PRO301 polypeptide. In particular, the 15 invention provides isolated native sequence PRO301 polypeptide, which in one embodiment, includes an amino acid sequence comprising the extracellular domain residues 28 to 258 of Figure 44 (SEQ ID NO:119). Native PRO301 polypeptides with or without the native signal sequence (amino acids 1 to 27 in Figure 44 (SEQ ID NO:119), and with or without the initiating methionine are specifically included. Additionally, the sequences 20 of the invention may also comprise the transmembrane domain (residues 236 to about 258 in Figure 44; SEQ ID NO:119) and/or the intracellular domain (about residue 259 to 299 in Figure 44; SEQ ID NO:119). Alternatively, the invention provides a PRO301 polypeptide encoded by the nucleic acid deposited under accession number ATCC 209432.

19. PRO224

25 Applicants have identified a cDNA clone that encodes a novel polypeptide, wherein the polypeptide is designated in the present application as "PRO224".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding 30 a PRO224 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO224 polypeptide having amino acid residues 1 to 282 of Figure 46 (SEQ ID NO:127), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO224 polypeptide. In particular, the invention provides isolated native sequence PRO224 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 282 of Figure 46 (SEQ ID NO:127).

35

20. PRO222

Applicants have identified a cDNA clone that encodes a novel polypeptide, wherein the polypeptide is

designated in the present application as "PRO222".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO222 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO222 polypeptide having amino acid residues 1 to 490 of Fig. 48 (SEQ ID NO:132), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under 5 high stringency conditions.

In another embodiment, the invention provides isolated PRO222 polypeptide. In particular, the invention provides isolated native sequence PRO222 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 490 of Figure 48 (SEQ ID NO:132).

10 21. PRO234

Applicants have identified a cDNA clone that encodes a novel lectin polypeptide molecule, designated in the present application as "PRO234".

In one embodiment, the invention provides an isolated nucleic acid encoding a novel lectin comprising DNA encoding a PRO234 polypeptide. In one aspect, the isolated nucleic acid comprises the DNA encoding 15 PRO234 polypeptides having amino acid residues 1 to 382 of Fig. 50 (SEQ ID NO:137), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions. In another aspect, the invention provides an isolated nucleic acid molecule comprising the nucleotide sequence of Fig. 49 (SEQ ID NO:136).

In another embodiment, the invention provides isolated novel PRO234 polypeptides. In particular, the 20 invention provides isolated native sequence PRO234 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 382 of Figure 50 (SEQ ID NO:137).

In yet another embodiment, the invention provides oligonucleotide probes useful for isolating genomic and cDNA nucleotide sequences.

25 22. PRO231

Applicants have identified a cDNA clone that encodes a novel polypeptide having homology to a putative acid phosphatase, wherein the polypeptide is designated in the present application as "PRO231".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO231 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO231 30 polypeptide having amino acid residues 1 to 428 of Fig. 52 (SEQ ID NO:142), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO231 polypeptide. In particular, the 35 invention provides isolated native sequence PRO231 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 428 of Figure 52 (SEQ ID NO:142).

23. PRO229

Applicants have identified a cDNA clone that encodes a novel polypeptide having homology to scavenger receptors wherein the polypeptide is designated in the present application as "PRO229".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO229 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO229 polypeptide having amino acid residues 1 to 347 of Figure 54 (SEQ ID NO:148), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO229 polypeptide. In particular, the invention provides isolated native sequence PRO229 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 347 of Figure 54 (SEQ ID NO:148).

24. PRO238

Applicants have identified a cDNA clone that encodes a novel polypeptide having homology to reductase, wherein the polypeptide is designated in the present application as "PRO238".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO238 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO238 polypeptide having amino acid residues 1 to 310 of Figure 56 (SEQ ID NO:153), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO238 polypeptide. In particular, the invention provides isolated native sequence PRO238 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 310 of Figure 56 (SEQ ID NO:153).

25. PRO233

Applicants have identified a cDNA clone that encodes a novel polypeptide, wherein the polypeptide is designated in the present application as "PRO233".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO233 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO233 polypeptide having amino acid residues 1 to 300 of Figure 58 (SEQ ID NO:159), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO233 polypeptide. In particular, the invention provides isolated native sequence PRO233 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 300 of Figure 58 (SEQ ID NO:159).

35

26. PRO223

Applicants have identified a cDNA clone that encodes a novel polypeptide having homology to serine

carboxypeptidase polypeptides, wherein the polypeptide is designated in the present application as "PRO223".

5 In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO223 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO223 polypeptide having amino acid residues 1 to 476 of Figure 60 (SEQ ID NO:164), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO223 polypeptide. In particular, the invention provides isolated native sequence PRO223 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 476 of Figure 60 (SEQ ID NO:164).

10 27. PRO235

Applicants have identified a cDNA clone that encodes a novel polypeptide, wherein the polypeptide is designated in the present application as "PRO235".

15 In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO235 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO235 polypeptide having amino acid residues 1 to 552 of Figure 62 (SEQ ID NO:170), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

20 In another embodiment, the invention provides isolated PRO235 polypeptide. In particular, the invention provides isolated native sequence PRO235 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 552 of Figure 62 (SEQ ID NO:170).

28. PRO236 and PRO262

Applicants have identified cDNA clones that encode novel polypeptides having homology to β -galactosidase, wherein those polypeptides are designated in the present application as "PRO236" and "PRO262".

25 In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO236 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO236 polypeptide having amino acid residues 1 to 636 of Figure 64 (SEQ ID NO:175), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

30 In another embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO262 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO262 polypeptide having amino acid residues 1 to 654 of Figure 66 (SEQ ID NO:177), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

35 In another embodiment, the invention provides isolated PRO236 polypeptide. In particular, the invention provides isolated native sequence PRO236 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 636 of Figure 64 (SEQ ID NO:175).

In another embodiment, the invention provides isolated PRO262 polypeptide. In particular, the invention provides isolated native sequence PRO262 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 654 of Figure 66 (SEQ ID NO:177).

29. PRO239

5 Applicants have identified a cDNA clone that encodes a novel polypeptide, wherein the polypeptide is designated in the present application as "PRO239".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO239 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO239 polypeptide having amino acid residues 1 to 501 of Figure 68 (SEQ ID NO:185), or is complementary to such 10 encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO239 polypeptide. In particular, the invention provides isolated native sequence PRO239 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 501 of Figure 68 (SEQ ID NO:185).

15

30. PRO257

Applicants have identified a cDNA clone that encodes a novel polypeptide, wherein the polypeptide is designated in the present application as "PRO257".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding 20 a PRO257 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO257 polypeptide having amino acid residues 1 to 607 of Figure 70 (SEQ ID NO:190), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO257 polypeptide. In particular, the 25 invention provides isolated native sequence PRO257 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 607 of Figure 70 (SEQ ID NO:190). An additional embodiment of the present invention is directed to an isolated extracellular domain of a PRO257 polypeptide.

31. PRO260

30 Applicants have identified a cDNA clone that encodes a novel polypeptide, wherein the polypeptide is designated in the present application as "PRO260".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding 35 a PRO260 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO260 polypeptide having amino acid residues 1 to 467 of Figure 72 (SEQ ID NO:195), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO260 polypeptide. In particular, the invention provides isolated native sequence PRO260 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 467 of Figure 72 (SEQ ID NO:195).

32. PRO263

5 Applicants have identified a cDNA clone that encodes a novel polypeptide having homology to CD44 antigen, wherein the polypeptide is designated in the present application as "PRO263".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO263 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO263 polypeptide having amino acid residues 1 to 322 of Figure 74 (SEQ ID NO:201), or is complementary to such 10 encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO263 polypeptide. In particular, the invention provides isolated native sequence PRO263 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 322 of Figure 74 (SEQ ID NO:201). An additional embodiment of the 15 present invention is directed to an isolated extracellular domain of a PRO263 polypeptide.

33. PRO270

Applicants have identified a cDNA clone that encodes a novel polypeptide, wherein the polypeptide is designated in the present application as "PRO270".

20 In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO270 polypeptide. In one aspect, the isolated nucleic acid comprises DNA which includes the sequence encoding the PRO270 polypeptide having amino acid residues 1 to 296 of Fig. 76 (SEQ ID NO:207), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

25 In another embodiment, the invention provides isolated PRO270 polypeptide. In particular, the invention provides isolated native sequence PRO270 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 296 of Figure 76 (SEQ ID NO:207).

34. PRO271

30 Applicants have identified a cDNA clone that encodes a novel polypeptide having homology to the proteoglycan link protein, wherein the polypeptide is designated in the present application as "PRO271".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO271 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO271 polypeptide having amino acid residues 1 to 360 of Figure 78 (SEQ ID NO:213), or is complementary to such 35 encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO271 polypeptide. In particular, the invention provides isolated native sequence PRO271 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 360 of Figure 78 (SEQ ID NO:213).

35. **PRO272**

5 Applicants have identified a cDNA clone that encodes a novel polypeptide, wherein the polypeptide is designated in the present application as "PRO272".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO272 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO272 polypeptide having amino acid residues 1 to 328 of Figure 80 (SEQ ID NO:221), or is complementary to such 10 encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally; under high stringency conditions.

In another embodiment, the invention provides isolated PRO272 polypeptide. In particular, the invention provides isolated native sequence PRO272 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 328 of Figure 80 (SEQ ID NO:211).

15

36. **PRO294**

Applicants have identified a cDNA clone that encodes a novel polypeptide, wherein the polypeptide is designated in the present application as "PRO294".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding 20 a PRO294 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO294 polypeptide having amino acid residues 1 to 550 of Figure 82 (SEQ ID NO:227), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO294 polypeptide. In particular, the 25 invention provides isolated native sequence PRO294 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 550 of Figure 82 (SEQ ID NO:227).

37. **PRO295**

Applicants have identified a cDNA clone that encodes a novel polypeptide, wherein the polypeptide is 30 designated in the present application as "PRO295".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO295 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO295 polypeptide having amino acid residues 1 to 350 of Figure 84 (SEQ ID NO:236), or is complementary to such 35 encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO295 polypeptide. In particular, the invention provides isolated native sequence PRO295 polypeptide, which in one embodiment, includes an amino

acid sequence comprising residues 1 to 350 of Figure 84 (SEQ ID NO:236).

38. **PRO293**

Applicants have identified a cDNA clone that encodes a novel human neuronal leucine rich repeat polypeptide, wherein the polypeptide is designated in the present application as "PRO293".

5 In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO293 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO293 polypeptide having amino acid residues 1 to 713 of Figure 86 (SEQ ID NO:245), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

10 In another embodiment, the invention provides isolated PRO293 polypeptide. In particular, the invention provides isolated native sequence PRO293 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 713 of Figure 86 (SEQ ID NO:245). An additional embodiment of the present invention is directed to an isolated extracellular domain of a PRO293 polypeptide.

15 39. **PRO247**

Applicants have identified a cDNA clone that encodes a novel polypeptide having leucine rich repeats wherein the polypeptide is designated in the present application as "PRO247".

20 In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO247 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO247 polypeptide having amino acid residues 1 to 546 of Figure 88 (SEQ ID NO:250), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

25 In another embodiment, the invention provides isolated PRO247 polypeptide. In particular, the invention provides isolated native sequence PRO247 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 546 of Figure 88 (SEQ ID NO:250). An additional embodiment of the present invention is directed to an isolated extracellular domain of a PRO247 polypeptide.

40. **PRO302, PRO303, PRO304, PRO307 and PRO343**

30 Applicants have identified cDNA clones that encode novel polypeptides having homology to various proteases, wherein those polypeptide are designated in the present application as "PRO302", "PRO303", "PRO304", "PRO307" and "PRO343" polypeptides.

35 In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO302 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO302 polypeptide having amino acid residues 1 to 452 of Figure 90 (SEQ ID NO:255), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO303 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO303 polypeptide having amino acid residues 1 to 314 of Figure 92 (SEQ ID NO:257), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

5 In yet another embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO304 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO304 polypeptide having amino acid residues 1 to 556 of Figure 94 (SEQ ID NO:259), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

10 In another embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO307 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO307 polypeptide having amino acid residues 1 to 383 of Figure 96 (SEQ ID NO:261), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

15 In another embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO343 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO343 polypeptide having amino acid residues 1 to 317 of Figure 98 (SEQ ID NO:263), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

20 In another embodiment, the invention provides isolated PRO302 polypeptide. In particular, the invention provides isolated native sequence PRO302 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 452 of Figure 90 (SEQ ID NO:255).

25 In another embodiment, the invention provides isolated PRO303 polypeptide. In particular, the invention provides isolated native sequence PRO303 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 314 of Figure 92 (SEQ ID NO:257).

In another embodiment, the invention provides isolated PRO304 polypeptide. In particular, the invention provides isolated native sequence PRO304 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 556 of Figure 94 (SEQ ID NO:259).

30 In another embodiment, the invention provides isolated PRO307 polypeptide. In particular, the invention provides isolated native sequence PRO307 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 383 of Figure 96 (SEQ ID NO:261).

In another embodiment, the invention provides isolated PRO343 polypeptide. In particular, the invention provides isolated native sequence PRO343 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 317 of Figure 98 (SEQ ID NO:263).

41. PRO328

Applicants have identified a cDNA clone that encodes a novel polypeptide, wherein the polypeptide is designated in the present application as "PRO328".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO328 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO328 polypeptide having amino acid residues 1 to 463 of Figure 100 (SEQ ID NO:285), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO328 polypeptide. In particular, the invention provides isolated native sequence PRO328 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 463 of Figure 100 (SEQ ID NO:285). An additional embodiment of the present invention is directed to an isolated extracellular domain of a PRO306 polypeptide.

42. PRO335, PRO331 and PRO326

Applicants have identified three cDNA clones that respectively encode three novel polypeptides, each 15 having leucine rich repeats and homology to LIG-1 and ALS. These polypeptides are designated in the present application as PRO335, PRO331 and PRO326, respectively.

In one embodiment, the invention provides three isolated nucleic acid molecules comprising DNA respectively encoding PRO335, PRO331 and PRO326, respectively. In one aspect, herein is provided an isolated nucleic acid comprising DNA encoding the PRO335 polypeptide having amino acid residues 1 through 20 1059 of Figure 102 (SEQ ID NO:290), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions. Also provided herein is an isolated nucleic acid comprises DNA encoding the PRO331 polypeptide having amino acid residues 1 through 640 of Figure 104 (SEQ ID NO:292), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

25 Additionally provided herein is an isolated nucleic acid comprises DNA encoding the PRO326 polypeptide having amino acid residues 1 through 1119 of Figure 106 (SEQ ID NO:294), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO335, PRO331 and PRO326 polypeptides 30 or extracellular domains thereof. In particular, the invention provides isolated native sequence for the PRO335 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 through 1059 of Figure 102 (SEQ ID NO:290). Also provided herein is the isolated native sequence for the PRO331 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 through 640 of Figure 104 (SEQ ID NO:292). Also provided herein is the isolated native sequence for the PRO326 polypeptide, 35 which in one embodiment, includes an amino acid sequence comprising residues 1 through 1119 of Figure 106 (SEQ ID NO:294).

43. PRO332

Applicants have identified a cDNA clone (DNA40982-1235) that encodes a novel polypeptide, designated in the present application as "PRO332."

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA having at least about 80% sequence identity to (a) a DNA molecule encoding a PRO358 polypeptide comprising the sequence of amino acids 49 to 642 of Fig. 108 (SEQ ID NO:310), or (b) the complement of the DNA molecule of (a). The sequence identity preferably is about 85%, more preferably about 90%, most preferably about 95%.

In one aspect, the isolated nucleic acid has at least about 80%, preferably at least about 85%, more preferably at least about 90%, and most preferably at least about 95% sequence identity with a polypeptide having amino acid residues 1 to 642 of Fig. 108 (SEQ ID NO:310). Preferably, the highest degree of sequence identity occurs within the leucine-rich repeat domains (amino acids 116 to 624 of Fig. 108, SEQ ID NO:310). In a further embodiment, the isolated nucleic acid molecule comprises DNA encoding a PRO332 polypeptide having amino acid residues 49 to 642 of Fig. 108 (SEQ ID NO:310), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

15 In another embodiment, the invention provides isolated PRO332 polypeptides. In particular, the invention provides isolated native sequence PRO332 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 49 to 624 of Figure 108 (SEQ ID NO:310). Native PRO332 polypeptides with or without the native signal sequence (amino acids 1 to 48 in Figure 108, SEQ ID NO:310), and with or without the initiating methionine are specifically included.

20

44. PRO334

Applicants have identified a cDNA clone that encodes a novel polypeptide having homology to fibulin and fibrillin, wherein the polypeptide is designated in the present application as "PRO334".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO334 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO334 polypeptide having amino acid residues 1 to 509 of Figure 110 (SEQ ID NO:315), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

30 In another embodiment, the invention provides isolated PRO334 polypeptide. In particular, the invention provides isolated native sequence PRO334 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 509 of Figure 110 (SEQ ID NO:315).

45. PRO346

Applicants have identified a cDNA clone (DNA44167-1243) that encodes a novel polypeptide, 35 designated in the present application as "PRO346."

In one embodiment, the invention provides an isolated nucleic acid molecule having at least about 80% sequence identity to (a) a DNA molecule encoding a PRO346 polypeptide comprising the sequence of amino

acids 19 to 339 of Fig. 112 (SEQ ID NO: 320), or (b) the complement of the DNA molecule of (a). The sequence identity preferably is about 85%, more preferably about 90%, most preferably about 95%. In one aspect, the isolated nucleic acid has at least about 80%, preferably at least about 85%, more preferably at least about 90%, and most preferably at least about 95% sequence identity with a polypeptide having amino acid residues 19 to 339 of Fig. 112 (SEQ ID NO:320). Preferably, the highest degree of sequence identity occurs 5 within the extracellular domains (amino acids 19 to 339 of Fig. 112, SEQ ID NO:320). In alternative embodiments, the polypeptide by which the homology is measured comprises the residues 1-339, 19-360 or 19-450 of Fig. 112, SEQ ID NO:320). In a further embodiment, the isolated nucleic acid molecule comprises DNA encoding a PRO346 polypeptide having amino acid residues 19 to 339 of Fig. 112 (SEQ ID NO:320), alternatively residues 1-339, 19-360 or 19-450 of Fig. 112 (SEQ ID NO:320) or is complementary to such 10 encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions. In another aspect, the invention provides a nucleic acid of the full length protein of clone DNA44167-1243, deposited with the ATCC under accession number ATCC 209434, alternatively the coding sequence of clone DNA44167-1243, deposited under accession number ATCC 209434.

In yet another embodiment, the invention provides isolated PRO346 polypeptide. In particular, the 15 invention provides isolated native sequence PRO346 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 19 to 339 of Figure 112 (SEQ ID NO:320). Native PRO346 polypeptides with or without the native signal sequence (residues 1 to 18 in Figure 112 (SEQ ID NO:320), with or without the initiating methionine, with or without the transmembrane domain (residues 340 to 360) and with or without the intracellular domain (residues 361 to 450) are specifically included. Alternatively, the invention provides 20 a PRO346 polypeptide encoded by the nucleic acid deposited under accession number ATCC 209434.

46. PRO268

Applicants have identified a cDNA clone that encodes a novel polypeptide having homology to protein disulfide isomerase, wherein the polypeptide is designated in the present application as "PRO268". 25 In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO268 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO268 polypeptide having amino acid residues 1 to 280 of Figure 114 (SEQ ID NO:325), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

30 In another embodiment, the invention provides isolated PRO268 polypeptide. In particular, the invention provides isolated native sequence PRO268 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 280 of Figure 114 (SEQ ID NO:325). An additional embodiment of the present invention is directed to an isolated extracellular domain of a PRO268 polypeptide.

35 47. PRO330

Applicants have identified a cDNA clone that encodes a novel polypeptide having homology to the alpha subunit of prolyl 4-hydroxylase, wherein the polypeptide is designated in the present application as "PRO330".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO330 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO330 polypeptide having amino acid residues 1 to 533 of Figure 116 (SEQ ID NO:332), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

5 In another embodiment, the invention provides isolated PRO330 polypeptide. In particular, the invention provides isolated native sequence PRO330 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 533 of Figure 116 (SEQ ID NO:332).

48. PRO339 and PRO310

10 Applicants have identified two cDNA clones wherein each clone encodes a novel polypeptide having homology to fringe, wherein the polypeptides are designated in the present application as "PRO339" and "PRO310".

15 In one embodiment, the invention provides isolated nucleic acid molecules comprising DNA encoding a PRO339 and/or a PRO310 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO339 polypeptide having amino acid residues 1 to 772 of Figure 118 (SEQ ID NO:339), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions. In another aspect, the isolated nucleic acid comprises DNA encoding the PRO310 polypeptide having amino acid residues 1 to 318 of Figure 120 (SEQ ID NO:341), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, 20 under high stringency conditions.

25 In another embodiment, the invention provides isolated PRO339 as well as isolated PRO310 polypeptides. In particular, the invention provides isolated native sequence PRO339 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 772 of Figure 118 (SEQ ID NO:339). The invention further provides isolated native sequence PRO310 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 318 of Figure 120 (SEQ ID NO:341).

49. PRO244

Applicants have identified a cDNA clone that encodes a novel polypeptide, designated in the present application as "PRO244".

30 In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding PRO244 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding PRO244 polypeptide having amino acid residues 1 to 219 of Fig. 122 (SEQ ID NO:377), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

35 In another embodiment, the invention provides isolated PRO244 polypeptide. In particular, the invention provides isolated native sequence PRO244 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 219 of Figure 122 (SEQ ID NO:377).

50. Additional Embodiments

In other embodiments of the present invention, the invention provides vectors comprising DNA encoding any of the herein described polypeptides. Host cell comprising any such vector are also provided. By way of example, the host cells may be CHO cells, *E. coli*, or yeast. A process for producing any of the herein described polypeptides is further provided and comprises culturing host cells under conditions suitable for expression of the desired polypeptide and recovering the desired polypeptide from the cell culture.

5 In other embodiments, the invention provides chimeric molecules comprising any of the herein described polypeptides fused to a heterologous polypeptide or amino acid sequence. Example of such chimeric molecules comprise any of the herein described polypeptides fused to an epitope tag sequence or a Fc region of an immunoglobulin.

10 In another embodiment, the invention provides an antibody which specifically binds to any of the above or below described polypeptides. Optionally, the antibody is a monoclonal antibody, humanized antibody, antibody fragment or single-chain antibody.

15 In yet other embodiments, the invention provides oligonucleotide probes useful for isolating genomic and cDNA nucleotide sequences, wherein those probes may be derived from any of the above or below described nucleotide sequences.

In other embodiments, the invention provides an isolated nucleic acid molecule comprising a nucleotide sequence that encodes a PRO polypeptide.

20 In one aspect, the isolated nucleic acid molecule comprises a nucleotide sequence having at least about 80% sequence identity, preferably at least about 81% sequence identity, more preferably at least about 82% sequence identity, yet more preferably at least about 83% sequence identity, yet more preferably at least about 84% sequence identity, yet more preferably at least about 85% sequence identity, yet more preferably at least about 86% sequence identity, yet more preferably at least about 87% sequence identity, yet more preferably at least about 88% sequence identity, yet more preferably at least about 89% sequence identity, yet more preferably at least about 90% sequence identity, yet more preferably at least about 91% sequence identity, yet more preferably at least about 92% sequence identity, yet more preferably at least about 93% sequence identity, yet more preferably at least about 94% sequence identity, yet more preferably at least about 95% sequence identity, yet more preferably at least about 96% sequence identity, yet more preferably at least about 97% sequence identity, yet more preferably at least about 98% sequence identity and yet more preferably at least about 99% sequence identity to (a) a DNA molecule encoding a PRO polypeptide having a full-length amino acid sequence as disclosed herein, an amino acid sequence lacking the signal peptide as disclosed herein or an extracellular domain of a transmembrane protein, with or without the signal peptide, as disclosed herein, or (b) the complement of the DNA molecule of (a).

30 In other aspects, the isolated nucleic acid molecule comprises a nucleotide sequence having at least about 80% sequence identity, preferably at least about 81% sequence identity, more preferably at least about 82% sequence identity, yet more preferably at least about 83% sequence identity, yet more preferably at least about 84% sequence identity, yet more preferably at least about 85% sequence identity, yet more preferably at least about 86% sequence identity, yet more preferably at least about 87% sequence identity, yet more preferably at

least about 88% sequence identity, yet more preferably at least about 89% sequence identity, yet more preferably at least about 90% sequence identity, yet more preferably at least about 91% sequence identity, yet more preferably at least about 92% sequence identity, yet more preferably at least about 93% sequence identity, yet more preferably at least about 94% sequence identity, yet more preferably at least about 95% sequence identity, yet more preferably at least about 96% sequence identity, yet more preferably at least about 97% sequence
5 identity, yet more preferably at least about 98% sequence identity and yet more preferably at least about 99% sequence identity to (a) a DNA molecule comprising the coding sequence of a full-length PRO polypeptide cDNA as disclosed herein, the coding sequence of a PRO polypeptide lacking the signal peptide as disclosed herein or the coding sequence of an extracellular domain of a transmembrane PRO polypeptide, with or without the signal peptide, as disclosed herein, or (b) the complement of the DNA molecule of (a).

10 In a further aspect, the invention concerns an isolated nucleic acid molecule comprising a nucleotide sequence having at least about 80% sequence identity, preferably at least about 81% sequence identity, more preferably at least about 82% sequence identity, yet more preferably at least about 83% sequence identity, yet more preferably at least about 84% sequence identity, yet more preferably at least about 85% sequence identity, yet more preferably at least about 86% sequence identity, yet more preferably at least about 87% sequence
15 identity, yet more preferably at least about 88% sequence identity, yet more preferably at least about 89% sequence identity, yet more preferably at least about 90% sequence identity, yet more preferably at least about 91% sequence identity, yet more preferably at least about 92% sequence identity, yet more preferably at least about 93% sequence identity, yet more preferably at least about 94% sequence identity, yet more preferably at least about 95% sequence identity, yet more preferably at least about 96% sequence identity, yet more preferably
20 at least about 97% sequence identity, yet more preferably at least about 98% sequence identity and yet more preferably at least about 99% sequence identity to (a) a DNA molecule that encodes the same mature polypeptide encoded by any of the human protein cDNAs deposited with the ATCC as disclosed herein, or (b) the complement of the DNA molecule of (a).

Another aspect the invention provides an isolated nucleic acid molecule comprising a nucleotide sequence encoding a PRO polypeptide which is either transmembrane domain-deleted or transmembrane domain-inactivated, or is complementary to such encoding nucleotide sequence, wherein the transmembrane domain(s) of such polypeptide are disclosed herein. Therefore, soluble extracellular domains of the herein described PRO polypeptides are contemplated.

Another embodiment is directed to fragments of a PRO polypeptide coding sequence, or the complement thereof, that may find use as, for example, hybridization probes or for encoding fragments of a PRO polypeptide that may optionally encode a polypeptide comprising a binding site for an anti-PRO antibody. Such nucleic acid fragments are usually at least about 20 nucleotides in length, preferably at least about 30 nucleotides in length, more preferably at least about 40 nucleotides in length, yet more preferably at least about 50 nucleotides in length, yet more preferably at least about 60 nucleotides in length, yet more preferably at least about 70 nucleotides in length, yet more preferably at least about 80 nucleotides in length, yet more preferably at least about 90 nucleotides in length, yet more preferably at least about 100 nucleotides in length, yet more preferably at least about 110 nucleotides in length, yet more preferably at least about 120 nucleotides in length, yet more

preferably at least about 130 nucleotides in length, yet more preferably at least about 140 nucleotides in length, yet more preferably at least about 150 nucleotides in length, yet more preferably at least about 160 nucleotides in length, yet more preferably at least about 170 nucleotides in length, yet more preferably at least about 180 nucleotides in length, yet more preferably at least about 190 nucleotides in length, yet more preferably at least about 200 nucleotides in length, yet more preferably at least about 250 nucleotides in length, yet more preferably at least about 300 nucleotides in length, yet more preferably at least about 350 nucleotides in length, yet more preferably at least about 400 nucleotides in length, yet more preferably at least about 450 nucleotides in length, yet more preferably at least about 500 nucleotides in length, yet more preferably at least about 600 nucleotides in length, yet more preferably at least about 700 nucleotides in length, yet more preferably at least about 800 nucleotides in length, yet more preferably at least about 900 nucleotides in length and yet more preferably at least about 1000 nucleotides in length, wherein in this context the term "about" means the referenced nucleotide sequence length plus or minus 10% of that referenced length. It is noted that novel fragments of a PRO polypeptide-encoding nucleotide sequence may be determined in a routine manner by aligning the PRO polypeptide-encoding nucleotide sequence with other known nucleotide sequences using any of a number of well known sequence alignment programs and determining which PRO polypeptide-encoding nucleotide sequence fragment(s) are novel. All of such PRO polypeptide-encoding nucleotide sequences are contemplated herein. Also contemplated are the PRO polypeptide fragments encoded by these nucleotide molecule fragments, preferably those PRO polypeptide fragments that comprise a binding site for an anti-PRO antibody.

In another embodiment, the invention provides isolated PRO polypeptide encoded by any of the isolated nucleic acid sequences hereinabove identified.

In a certain aspect, the invention concerns an isolated PRO polypeptide, comprising an amino acid sequence having at least about 80% sequence identity, preferably at least about 81% sequence identity, more preferably at least about 82% sequence identity, yet more preferably at least about 83% sequence identity, yet more preferably at least about 84% sequence identity, yet more preferably at least about 85% sequence identity, yet more preferably at least about 86% sequence identity, yet more preferably at least about 87% sequence identity, yet more preferably at least about 88% sequence identity, yet more preferably at least about 89% sequence identity, yet more preferably at least about 90% sequence identity, yet more preferably at least about 91% sequence identity, yet more preferably at least about 92% sequence identity, yet more preferably at least about 93% sequence identity, yet more preferably at least about 94% sequence identity, yet more preferably at least about 95% sequence identity, yet more preferably at least about 96% sequence identity, yet more preferably at least about 97% sequence identity, yet more preferably at least about 98% sequence identity and yet more preferably at least about 99% sequence identity to a PRO polypeptide having a full-length amino acid sequence as disclosed herein, an amino acid sequence lacking the signal peptide as disclosed herein or an extracellular domain of a transmembrane protein, with or without the signal peptide, as disclosed herein.

In a further aspect, the invention concerns an isolated PRO polypeptide comprising an amino acid sequence having at least about 80% sequence identity, preferably at least about 81% sequence identity, more preferably at least about 82% sequence identity, yet more preferably at least about 83% sequence identity, yet more preferably at least about 84% sequence identity, yet more preferably at least about 85% sequence identity,

yet more preferably at least about 86% sequence identity, yet more preferably at least about 87% sequence identity, yet more preferably at least about 88% sequence identity, yet more preferably at least about 89% sequence identity, yet more preferably at least about 90% sequence identity, yet more preferably at least about 91% sequence identity, yet more preferably at least about 92% sequence identity, yet more preferably at least about 93% sequence identity, yet more preferably at least about 94% sequence identity, yet more preferably at least about 95% sequence identity, yet more preferably at least about 96% sequence identity, yet more preferably at least about 97% sequence identity, yet more preferably at least about 98% sequence identity and yet more preferably at least about 99% sequence identity to an amino acid sequence encoded by any of the human protein cDNAs deposited with the ATCC as disclosed herein.

In a further aspect, the invention concerns an isolated PRO polypeptide comprising an amino acid sequence scoring at least about 80% positives, preferably at least about 81% positives, more preferably at least about 82% positives, yet more preferably at least about 83% positives, yet more preferably at least about 84% positives, yet more preferably at least about 85% positives, yet more preferably at least about 86% positives, yet more preferably at least about 87% positives, yet more preferably at least about 88% positives, yet more preferably at least about 89% positives, yet more preferably at least about 90% positives, yet more preferably at least about 91% positives, yet more preferably at least about 92% positives, yet more preferably at least about 93% positives, yet more preferably at least about 94% positives, yet more preferably at least about 95% positives, yet more preferably at least about 96% positives, yet more preferably at least about 97% positives, yet more preferably at least about 98% positives and yet more preferably at least about 99% positives when compared with the amino acid sequence of a PRO polypeptide having a full-length amino acid sequence as disclosed herein, an amino acid sequence lacking the signal peptide as disclosed herein or an extracellular domain of a transmembrane protein, with or without the signal peptide, as disclosed herein.

In a specific aspect, the invention provides an isolated PRO polypeptide without the N-terminal signal sequence and/or the initiating methionine and is encoded by a nucleotide sequence that encodes such an amino acid sequence as hereinbefore described. Processes for producing the same are also herein described, wherein those processes comprise culturing a host cell comprising a vector which comprises the appropriate encoding nucleic acid molecule under conditions suitable for expression of the PRO polypeptide and recovering the PRO polypeptide from the cell culture.

Another aspect the invention provides an isolated PRO polypeptide which is either transmembrane domain-deleted or transmembrane domain-inactivated. Processes for producing the same are also herein described, wherein those processes comprise culturing a host cell comprising a vector which comprises the appropriate encoding nucleic acid molecule under conditions suitable for expression of the PRO polypeptide and recovering the PRO polypeptide from the cell culture.

In yet another embodiment, the invention concerns agonists and antagonists of a native PRO polypeptide as defined herein. In a particular embodiment, the agonist or antagonist is an anti-PRO antibody or a small molecule.

In a further embodiment, the invention concerns a method of identifying agonists or antagonists to a PRO polypeptide which comprise contacting the PRO polypeptide with a candidate molecule and monitoring a

biological activity mediated by said PRO polypeptide. Preferably, the PRO polypeptide is a native PRO polypeptide.

In a still further embodiment, the invention concerns a composition of matter comprising a PRO polypeptide, or an agonist or antagonist of a PRO polypeptide as herein described, or an anti-PRO antibody, in combination with a carrier. Optionally, the carrier is a pharmaceutically acceptable carrier.

5 Another embodiment of the present invention is directed to the use of a PRO polypeptide, or an agonist or antagonist thereof as hereinbefore described, or an anti-PRO antibody, for the preparation of a medicament useful in the treatment of a condition which is responsive to the PRO polypeptide, an agonist or antagonist thereof or an anti-PRO antibody.

10

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 shows a nucleotide sequence (SEQ ID NO:1) of a native sequence PRO211 cDNA, wherein SEQ ID NO:1 is a clone designated herein as "DNA32292-1131".

Figure 2 shows the amino acid sequence (SEQ ID NO:2) derived from the coding sequence of SEQ ID NO:1 shown in Figure 1.

15

Figure 3 shows a nucleotide sequence (SEQ ID NO:3) of a native sequence PRO217 cDNA, wherein SEQ ID NO:3 is a clone designated herein as "DNA33094-1131".

Figure 4 shows the amino acid sequence (SEQ ID NO:4) derived from the coding sequence of SEQ ID NO:3 shown in Figure 3.

20

Figure 5 shows a nucleotide sequence (SEQ ID NO:11) of a native sequence PRO230 cDNA, wherein SEQ ID NO:11 is a clone designated herein as "DNA33223-1136".

Figure 6 shows the amino acid sequence (SEQ ID NO:12) derived from the coding sequence of SEQ ID NO:11 shown in Figure 5.

Figure 7 shows a nucleotide sequence designated herein as DNA20088 (SEQ ID NO:13).

25

Figure 8 shows a nucleotide sequence (SEQ ID NO:17) of a native sequence PRO232 cDNA, wherein SEQ ID NO:17 is a clone designated herein as "DNA34435-1140".

Figure 9 shows the amino acid sequence (SEQ ID NO:18) derived from the coding sequence of SEQ ID NO:17 shown in Figure 8.

Figure 10 shows a nucleotide sequence (SEQ ID NO:22) of a native sequence PRO187 cDNA, wherein SEQ ID NO:22 is a clone designated herein as "DNA27864-1155".

30

Figure 11 shows the amino acid sequence (SEQ ID NO:23) derived from the coding sequence of SEQ ID NO:22 shown in Figure 10.

Figure 12 shows a nucleotide sequence (SEQ ID NO:27) of a native sequence PRO265 cDNA, wherein SEQ ID NO:27 is a clone designated herein as "DNA36350-1158".

35

Figure 13 shows the amino acid sequence (SEQ ID NO:28) derived from the coding sequence of SEQ ID NO:27 shown in Figure 12.

Figure 14 shows a nucleotide sequence (SEQ ID NO:33) of a native sequence PRO219 cDNA, wherein SEQ ID NO:33 is a clone designated herein as "DNA32290-1164".

Figure 15 shows the amino acid sequence (SEQ ID NO:34) derived from the coding sequence of SEQ ID NO:33 shown in Figure 14.

Figure 16 shows a nucleotide sequence (SEQ ID NO:38) of a native sequence PRO246 cDNA, wherein SEQ ID NO:38 is a clone designated herein as "DNA35639-1172".

5 Figure 17 shows the amino acid sequence (SEQ ID NO:39) derived from the coding sequence of SEQ ID NO:38 shown in Figure 16.

Figure 18 shows a nucleotide sequence (SEQ ID NO:48) of a native sequence PRO228 cDNA, wherein SEQ ID NO:48 is a clone designated herein as "DNA33092-1202".

Figure 19 shows the amino acid sequence (SEQ ID NO:49) derived from the coding sequence of SEQ ID NO:48 shown in Figure 18.

10 Figure 20 shows a nucleotide sequence designated herein as DNA21951 (SEQ ID NO:50).

Figure 21 shows a nucleotide sequence (SEQ ID NO:58) of a native sequence PRO533 cDNA, wherein SEQ ID NO:58 is a clone designated herein as "DNA49435-1219".

Figure 22 shows the amino acid sequence (SEQ ID NO:59) derived from the coding sequence of SEQ ID NO:58 shown in Figure 21.

15 Figure 23 shows a nucleotide sequence (SEQ ID NO:63) of a native sequence PRO245 cDNA, wherein SEQ ID NO:63 is a clone designated herein as "DNA35638-1141".

Figure 24 shows the amino acid sequence (SEQ ID NO:64) derived from the coding sequence of SEQ ID NO:63 shown in Figure 23.

20 Figure 25 shows a nucleotide sequence (SEQ ID NO:68) of a native sequence PRO220 cDNA, wherein SEQ ID NO:68 is a clone designated herein as "DNA32298-1132".

Figure 26 shows the amino acid sequence (SEQ ID NO:69) derived from the coding sequence of SEQ ID NO:68 shown in Figure 25.

Figure 27 shows a nucleotide sequence (SEQ ID NO:70) of a native sequence PRO221 cDNA, wherein SEQ ID NO:70 is a clone designated herein as "DNA33089-1132".

25 Figure 28 shows the amino acid sequence (SEQ ID NO:71) derived from the coding sequence of SEQ ID NO:70 shown in Figure 27.

Figure 29 shows a nucleotide sequence (SEQ ID NO:72) of a native sequence PRO227 cDNA, wherein SEQ ID NO:72 is a clone designated herein as "DNA33786-1132".

30 Figure 30 shows the amino acid sequence (SEQ ID NO:73) derived from the coding sequence of SEQ ID NO:72 shown in Figure 29.

Figure 31 shows a nucleotide sequence (SEQ ID NO:83) of a native sequence PRO258 cDNA, wherein SEQ ID NO:83 is a clone designated herein as "DNA35918-1174".

Figure 32 shows the amino acid sequence (SEQ ID NO:84) derived from the coding sequence of SEQ ID NO:83 shown in Figure 31.

35 Figure 33 shows a nucleotide sequence (SEQ ID NO:90) of a native sequence PRO266 cDNA, wherein SEQ ID NO:90 is a clone designated herein as "DNA37150-1178".

Figure 34 shows the amino acid sequence (SEQ ID NO:91) derived from the coding sequence of SEQ ID NO:90 shown in Figure 33.

Figure 35 shows a nucleotide sequence (SEQ ID NO:95) of a native sequence PRO269 cDNA, wherein SEQ ID NO:95 is a clone designated herein as "DNA38260-1180".

5 Figure 36 shows the amino acid sequence (SEQ ID NO:96) derived from the coding sequence of SEQ ID NO:95 shown in Figure 35.

Figure 37 shows a nucleotide sequence (SEQ ID NO:103) of a native sequence PRO287 cDNA, wherein SEQ ID NO:103 is a clone designated herein as "DNA39969-1185".

Figure 38 shows the amino acid sequence (SEQ ID NO:104) derived from the coding sequence of SEQ ID NO:103 shown in Figure 37.

10 Figure 39 shows a nucleotide sequence (SEQ ID NO:108) of a native sequence PRO214 cDNA, wherein SEQ ID NO:108 is a clone designated herein as "DNA32286-1191".

Figure 40 shows the amino acid sequence (SEQ ID NO:109) derived from the coding sequence of SEQ ID NO:108 shown in Figure 39.

15 Figure 41 shows a nucleotide sequence (SEQ ID NO:113) of a native sequence PRO317 cDNA, wherein SEQ ID NO:113 is a clone designated herein as "DNA33461-1199".

Figure 42 shows the amino acid sequence (SEQ ID NO:114) derived from the coding sequence of SEQ ID NO:113 shown in Figure 41.

Figure 43 shows a nucleotide sequence (SEQ ID NO:118) of a native sequence PRO301 cDNA, wherein SEQ ID NO:118 is a clone designated herein as "DNA40628-1216".

20 Figure 44 shows the amino acid sequence (SEQ ID NO:119) derived from the coding sequence of SEQ ID NO:118 shown in Figure 43.

Figure 45 shows a nucleotide sequence (SEQ ID NO:126) of a native sequence PRO224 cDNA, wherein SEQ ID NO:126 is a clone designated herein as "DNA33221-1133".

25 Figure 46 shows the amino acid sequence (SEQ ID NO:127) derived from the coding sequence of SEQ ID NO:126 shown in Figure 45.

Figure 47 shows a nucleotide sequence (SEQ ID NO:131) of a native sequence PRO222 cDNA, wherein SEQ ID NO:131 is a clone designated herein as "DNA33107-1135".

Figure 48 shows the amino acid sequence (SEQ ID NO:132) derived from the coding sequence of SEQ ID NO:131 shown in Figure 47.

30 Figure 49 shows a nucleotide sequence (SEQ ID NO:136) of a native sequence PRO234 cDNA, wherein SEQ ID NO:136 is a clone designated herein as "DNA35557-1137".

Figure 50 shows the amino acid sequence (SEQ ID NO:137) derived from the coding sequence of SEQ ID NO:136 shown in Figure 49.

35 Figure 51 shows a nucleotide sequence (SEQ ID NO:141) of a native sequence PRO231 cDNA, wherein SEQ ID NO:141 is a clone designated herein as "DNA34434-1139".

Figure 52 shows the amino acid sequence (SEQ ID NO:142) derived from the coding sequence of SEQ ID NO:141 shown in Figure 51.

Figure 53 shows a nucleotide sequence (SEQ ID NO:147) of a native sequence PRO229 cDNA, wherein SEQ ID NO:147 is a clone designated herein as "DNA33100-1159".

Figure 54 shows the amino acid sequence (SEQ ID NO:148) derived from the coding sequence of SEQ ID NO:147 shown in Figure 53.

5 Figure 55 shows a nucleotide sequence (SEQ ID NO:152) of a native sequence PRO238 cDNA, wherein SEQ ID NO:152 is a clone designated herein as "DNA35600-1162".

Figure 56 shows the amino acid sequence (SEQ ID NO:153) derived from the coding sequence of SEQ ID NO:152 shown in Figure 55.

Figure 57 shows a nucleotide sequence (SEQ ID NO:158) of a native sequence PRO233 cDNA, wherein SEQ ID NO:158 is a clone designated herein as "DNA34436-1238".

10 Figure 58 shows the amino acid sequence (SEQ ID NO:159) derived from the coding sequence of SEQ ID NO:158 shown in Figure 57.

Figure 59 shows a nucleotide sequence (SEQ ID NO:163) of a native sequence PRO223 cDNA, wherein SEQ ID NO:163 is a clone designated herein as "DNA33206-1165".

15 Figure 60 shows the amino acid sequence (SEQ ID NO:164) derived from the coding sequence of SEQ ID NO:163 shown in Figure 59.

Figure 61 shows a nucleotide sequence (SEQ ID NO:169) of a native sequence PRO235 cDNA, wherein SEQ ID NO:169 is a clone designated herein as "DNA35558-1167".

Figure 62 shows the amino acid sequence (SEQ ID NO:170) derived from the coding sequence of SEQ ID NO:169 shown in Figure 61.

20 Figure 63 shows a nucleotide sequence (SEQ ID NO:174) of a native sequence PRO236 cDNA, wherein SEQ ID NO:174 is a clone designated herein as "DNA35599-1168".

Figure 64 shows the amino acid sequence (SEQ ID NO:175) derived from the coding sequence of SEQ ID NO:174 shown in Figure 63.

25 Figure 65 shows a nucleotide sequence (SEQ ID NO:176) of a native sequence PRO262 cDNA, wherein SEQ ID NO:176 is a clone designated herein as "DNA36992-1168".

Figure 66 shows the amino acid sequence (SEQ ID NO:177) derived from the coding sequence of SEQ ID NO:176 shown in Figure 65.

Figure 67 shows a nucleotide sequence (SEQ ID NO:184) of a native sequence PRO239 cDNA, wherein SEQ ID NO:184 is a clone designated herein as "DNA34407-1169".

30 Figure 68 shows the amino acid sequence (SEQ ID NO:185) derived from the coding sequence of SEQ ID NO:184 shown in Figure 67.

Figure 69 shows a nucleotide sequence (SEQ ID NO:189) of a native sequence PRO257 cDNA, wherein SEQ ID NO:189 is a clone designated herein as "DNA35841-1173".

35 Figure 70 shows the amino acid sequence (SEQ ID NO:190) derived from the coding sequence of SEQ ID NO:189 shown in Figure 69.

Figure 71 shows a nucleotide sequence (SEQ ID NO:194) of a native sequence PRO260 cDNA, wherein SEQ ID NO:194 is a clone designated herein as "DNA33470-1175".

Figure 72 shows the amino acid sequence (SEQ ID NO:195) derived from the coding sequence of SEQ ID NO:194 shown in Figure 71.

Figure 73 shows a nucleotide sequence (SEQ ID NO:200) of a native sequence PRO263 cDNA, wherein SEQ ID NO:200 is a clone designated herein as "DNA34431-1177".

Figure 74 shows the amino acid sequence (SEQ ID NO:201) derived from the coding sequence of SEQ ID NO:200 shown in Figure 73.

Figure 75 shows a nucleotide sequence (SEQ ID NO:206) of a native sequence PRO270 cDNA, wherein SEQ ID NO:206 is a clone designated herein as "DNA39510-1181".

Figure 76 shows the amino acid sequence (SEQ ID NO:207) derived from the coding sequence of SEQ ID NO:206 shown in Figure 75.

Figure 77 shows a nucleotide sequence (SEQ ID NO:212) of a native sequence PRO271 cDNA, wherein SEQ ID NO:212 is a clone designated herein as "DNA39423-1182".

Figure 78 shows the amino acid sequence (SEQ ID NO:213) derived from the coding sequence of SEQ ID NO:212 shown in Figure 77.

Figure 79 shows a nucleotide sequence (SEQ ID NO:220) of a native sequence PRO272 cDNA, wherein SEQ ID NO:220 is a clone designated herein as "DNA40620-1183".

Figure 80 shows the amino acid sequence (SEQ ID NO:221) derived from the coding sequence of SEQ ID NO:220 shown in Figure 79.

Figure 81 shows a nucleotide sequence (SEQ ID NO:226) of a native sequence PRO294 cDNA, wherein SEQ ID NO:226 is a clone designated herein as "DNA40604-1187".

Figure 82 shows the amino acid sequence (SEQ ID NO:227) derived from the coding sequence of SEQ ID NO:226 shown in Figure 81.

Figure 83 shows a nucleotide sequence (SEQ ID NO:235) of a native sequence PRO295 cDNA, wherein SEQ ID NO:235 is a clone designated herein as "DNA38268-1188".

Figure 84 shows the amino acid sequence (SEQ ID NO:236) derived from the coding sequence of SEQ ID NO:235 shown in Figure 83.

Figure 85 shows a nucleotide sequence (SEQ ID NO:244) of a native sequence PRO293 cDNA, wherein SEQ ID NO:244 is a clone designated herein as "DNA37151-1193".

Figure 86 shows the amino acid sequence (SEQ ID NO:245) derived from the coding sequence of SEQ ID NO:244 shown in Figure 85.

Figure 87 shows a nucleotide sequence (SEQ ID NO:249) of a native sequence PRO247 cDNA, wherein SEQ ID NO:249 is a clone designated herein as "DNA35673-1201".

Figure 88 shows the amino acid sequence (SEQ ID NO:250) derived from the coding sequence of SEQ ID NO:249 shown in Figure 87.

Figure 89 shows a nucleotide sequence (SEQ ID NO:254) of a native sequence PRO302 cDNA, wherein SEQ ID NO:254 is a clone designated herein as "DNA40370-1217".

Figure 90 shows the amino acid sequence (SEQ ID NO:255) derived from the coding sequence of SEQ ID NO:254 shown in Figure 89.

Figure 91 shows a nucleotide sequence (SEQ ID NO:256) of a native sequence PRO303 cDNA, wherein SEQ ID NO:256 is a clone designated herein as "DNA42551-1217".

Figure 92 shows the amino acid sequence (SEQ ID NO:257) derived from the coding sequence of SEQ ID NO:256 shown in Figure 91.

Figure 93 shows a nucleotide sequence (SEQ ID NO:258) of a native sequence PRO304 cDNA, wherein

5 SEQ ID NO:258 is a clone designated herein as "DNA39520-1217".

Figure 94 shows the amino acid sequence (SEQ ID NO:259) derived from the coding sequence of SEQ ID NO:258 shown in Figure 93.

Figure 95 shows a nucleotide sequence (SEQ ID NO:260) of a native sequence PRO307 cDNA, wherein SEQ ID NO:260 is a clone designated herein as "DNA41225-1217".

10 Figure 96 shows the amino acid sequence (SEQ ID NO:261) derived from the coding sequence of SEQ ID NO:260 shown in Figure 95.

Figure 97 shows a nucleotide sequence (SEQ ID NO:262) of a native sequence PRO343 cDNA, wherein SEQ ID NO:262 is a clone designated herein as "DNA43318-1217".

15 Figure 98 shows the amino acid sequence (SEQ ID NO:263) derived from the coding sequence of SEQ ID NO:262 shown in Figure 97.

Figure 99 shows a nucleotide sequence (SEQ ID NO:284) of a native sequence PRO328 cDNA, wherein SEQ ID NO:284 is a clone designated herein as "DNA40587-1231".

Figure 100 shows the amino acid sequence (SEQ ID NO:285) derived from the coding sequence of SEQ ID NO:284 shown in Figure 99.

20 Figure 101 shows a nucleotide sequence (SEQ ID NO:289) of a native sequence PRO335 cDNA, wherein SEQ ID NO:289 is a clone designated herein as "DNA41388-1234".

Figure 102 shows the amino acid sequence (SEQ ID NO:290) derived from the coding sequence of SEQ ID NO:289 shown in Figure 101.

25 Figure 103 shows a nucleotide sequence (SEQ ID NO:291) of a native sequence PRO331 cDNA, wherein SEQ ID NO:291 is a clone designated herein as "DNA40981-1234".

Figure 104 shows the amino acid sequence (SEQ ID NO:292) derived from the coding sequence of SEQ ID NO:291 shown in Figure 103.

Figure 105 shows a nucleotide sequence (SEQ ID NO:293) of a native sequence PRO326 cDNA, wherein SEQ ID NO:293 is a clone designated herein as "DNA37140-1234".

30 Figure 106 shows the amino acid sequence (SEQ ID NO:294) derived from the coding sequence of SEQ ID NO:293 shown in Figure 105.

Figure 107 shows a nucleotide sequence (SEQ ID NO:309) of a native sequence PRO332 cDNA, wherein SEQ ID NO:309 is a clone designated herein as "DNA40982-1235".

35 Figure 108 shows the amino acid sequence (SEQ ID NO:310) derived from the coding sequence of SEQ ID NO:309 shown in Figure 107.

Figure 109 shows a nucleotide sequence (SEQ ID NO:314) of a native sequence PRO334 cDNA, wherein SEQ ID NO:314 is a clone designated herein as "DNA41379-1236".

Figure 110 shows the amino acid sequence (SEQ ID NO:315) derived from the coding sequence of SEQ ID NO:314 shown in Figure 109.

Figure 111 shows a nucleotide sequence (SEQ ID NO:319) of a native sequence PRO346 cDNA, wherein SEQ ID NO:319 is a clone designated herein as "DNA44167-1243".

Figure 112 shows the amino acid sequence (SEQ ID NO:320) derived from the coding sequence of SEQ 5 ID NO:319 shown in Figure 111.

Figure 113 shows a nucleotide sequence (SEQ ID NO:324) of a native sequence PRO268 cDNA, wherein SEQ ID NO:324 is a clone designated herein as "DNA39427-1179".

Figure 114 shows the amino acid sequence (SEQ ID NO:325) derived from the coding sequence of SEQ ID NO:324 shown in Figure 113.

10 Figure 115 shows a nucleotide sequence (SEQ ID NO:331) of a native sequence PRO330 cDNA, wherein SEQ ID NO:331 is a clone designated herein as "DNA40603-1232".

Figure 116 shows the amino acid sequence (SEQ ID NO:332) derived from the coding sequence of SEQ ID NO:331 shown in Figure 115.

15 Figure 117 shows a nucleotide sequence (SEQ ID NO:338) of a native sequence PRO339 cDNA, wherein SEQ ID NO:338 is a clone designated herein as "DNA43466-1225".

Figure 118 shows the amino acid sequence (SEQ ID NO:339) derived from the coding sequence of SEQ ID NO:338 shown in Figure 117.

Figure 119 shows a nucleotide sequence (SEQ ID NO:340) of a native sequence PRO310 cDNA, wherein SEQ ID NO:340 is a clone designated herein as "DNA43046-1225".

20 Figure 120 shows the amino acid sequence (SEQ ID NO:341) derived from the coding sequence of SEQ ID NO:340 shown in Figure 119.

Figure 121 shows a nucleotide sequence (SEQ ID NO:376) of a native sequence PRO244 cDNA, wherein SEQ ID NO:376 is a clone designated herein as "DNA35668-1171".

25 Figure 122 shows the amino acid sequence (SEQ ID NO:377) derived from the coding sequence of SEQ ID NO:376 shown in Figure 121.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTSI. Definitions

The terms "PRO polypeptide" and "PRO" as used herein and when immediately followed by a numerical designation refer to various polypeptides, wherein the complete designation (i.e., PRO/number) refers to specific polypeptide sequences as described herein. The terms "PRO/number polypeptide" and 5 "PRO/number" wherein the term "number" is provided as an actual numerical designation as used herein encompass native sequence polypeptides and polypeptide variants (which are further defined herein). The PRO polypeptides described herein may be isolated from a variety of sources, such as from human tissue types or from another source, or prepared by recombinant or synthetic methods.

A "native sequence PRO polypeptide" comprises a polypeptide having the same amino acid sequence 10 as the corresponding PRO polypeptide derived from nature. Such native sequence PRO polypeptides can be isolated from nature or can be produced by recombinant or synthetic means. The term "native sequence PRO polypeptide" specifically encompasses naturally-occurring truncated or secreted forms of the specific PRO polypeptide (e.g., an extracellular domain sequence), naturally-occurring variant forms (e.g., alternatively spliced forms) and naturally-occurring allelic variants of the polypeptide. In various embodiments of the 15 invention, the native sequence PRO polypeptides disclosed herein are mature or full-length native sequence polypeptides comprising the full-length amino acids sequences shown in the accompanying figures. Start and stop codons are shown in bold font and underlined in the figures. However, while the PRO polypeptide disclosed in the accompanying figures are shown to begin with methionine residues designated herein as amino acid position 1 in the figures, it is conceivable and possible that other methionine residues located either upstream 20 or downstream from the amino acid position 1 in the figures may be employed as the starting amino acid residue for the PRO polypeptides.

The PRO polypeptide "extracellular domain" or "ECD" refers to a form of the PRO polypeptide which is essentially free of the transmembrane and cytoplasmic domains. Ordinarily, a PRO polypeptide ECD will have less than 1% of such transmembrane and/or cytoplasmic domains and preferably, will have less than 0.5% of 25 such domains. It will be understood that any transmembrane domains identified for the PRO polypeptides of the present invention are identified pursuant to criteria routinely employed in the art for identifying that type of hydrophobic domain. The exact boundaries of a transmembrane domain may vary but most likely by no more than about 5 amino acids at either end of the domain as initially identified herein. Optionally, therefore, an extracellular domain of a PRO polypeptide may contain from about 5 or fewer amino acids on either side of the 30 transmembrane domain/extracellular domain boundary as identified in the Examples or specification and such polypeptides, with or without the associated signal peptide, and nucleic acid encoding them, are contemplated by the present invention.

The approximate location of the "signal peptides" of the various PRO polypeptides disclosed herein are shown in the accompanying figures. It is noted, however, that the C-terminal boundary of a signal peptide may 35 vary, but most likely by no more than about 5 amino acids on either side of the signal peptide C-terminal boundary as initially identified herein, wherein the C-terminal boundary of the signal peptide may be identified pursuant to criteria routinely employed in the art for identifying that type of amino acid sequence element (e.g.,

Nielsen et al., Prot. Eng. 10:1-6 (1997) and von Heinje et al., Nucl. Acids. Res. 14:4683-4690 (1986)). Moreover, it is also recognized that, in some cases, cleavage of a signal sequence from a secreted polypeptide is not entirely uniform, resulting in more than one secreted species. These mature polypeptides, where the signal peptide is cleaved within no more than about 5 amino acids on either side of the C-terminal boundary of the signal peptide as identified herein, and the polynucleotides encoding them, are contemplated by the present 5 invention.

"PRO polypeptide variant" means an active PRO polypeptide as defined above or below having at least about 80% amino acid sequence identity with a full-length native sequence PRO polypeptide sequence as disclosed herein, a PRO polypeptide sequence lacking the signal peptide as disclosed herein, an extracellular domain of a PRO polypeptide, with or without the signal peptide, as disclosed herein or any other fragment of 10 a full-length PRO polypeptide sequence as disclosed herein. Such PRO polypeptide variants include, for instance, PRO polypeptides wherein one or more amino acid residues are added, or deleted, at the N- or C- terminus of the full-length native amino acid sequence. Ordinarily, a PRO polypeptide variant will have at least about 80% amino acid sequence identity, preferably at least about 81% amino acid sequence identity, more preferably at least about 82% amino acid sequence identity, more preferably at least about 83% amino acid 15 sequence identity, more preferably at least about 84% amino acid sequence identity, more preferably at least about 85% amino acid sequence identity, more preferably at least about 86% amino acid sequence identity, more preferably at least about 87% amino acid sequence identity, more preferably at least about 88% amino acid sequence identity, more preferably at least about 89% amino acid sequence identity, more preferably at least about 90% amino acid sequence identity, more preferably at least about 91% amino acid sequence identity, more 20 preferably at least about 92% amino acid sequence identity, more preferably at least about 93% amino acid sequence identity, more preferably at least about 94% amino acid sequence identity, more preferably at least about 95% amino acid sequence identity, more preferably at least about 96% amino acid sequence identity, more preferably at least about 97% amino acid sequence identity, more preferably at least about 98% amino acid sequence identity and most preferably at least about 99% amino acid sequence identity with a full-length native 25 sequence PRO polypeptide sequence as disclosed herein, a PRO polypeptide sequence lacking the signal peptide as disclosed herein, an extracellular domain of a PRO polypeptide, with or without the signal peptide, as disclosed herein or any other fragment of a full-length PRO polypeptide sequence as disclosed herein. Ordinarily, PRO variant polypeptides are at least about 10 amino acids in length, often at least about 20 amino acids in length, more often at least about 30 amino acids in length, more often at least about 40 amino acids in 30 length, more often at least about 50 amino acids in length, more often at least about 60 amino acids in length, more often at least about 70 amino acids in length, more often at least about 80 amino acids in length, more often at least about 90 amino acids in length, more often at least about 100 amino acids in length, more often at least about 150 amino acids in length, more often at least about 200 amino acids in length, more often at least about 300 amino acids in length, or more.

35 "Percent (%) amino acid sequence identity" with respect to the PRO polypeptide sequences identified herein is defined as the percentage of amino acid residues in a candidate sequence that are identical with the amino acid residues in the specific PRO polypeptide sequence, after aligning the sequences and introducing gaps,

if necessary, to achieve the maximum percent sequence identity, and not considering any conservative substitutions as part of the sequence identity. Alignment for purposes of determining percent amino acid sequence identity can be achieved in various ways that are within the skill in the art, for instance, using publicly available computer software such as BLAST, BLAST-2, ALIGN or Megalign (DNASTAR) software. Those skilled in the art can determine appropriate parameters for measuring alignment, including any algorithms needed 5 to achieve maximal alignment over the full length of the sequences being compared. For purposes herein, however, % amino acid sequence identity values are generated using the sequence comparison computer program ALIGN-2, wherein the complete source code for the ALIGN-2 program is provided in Table 10 below. The ALIGN-2 sequence comparison computer program was authored by Genentech, Inc. and the source code shown in Table 10 below has been filed with user documentation in the U.S. Copyright Office, Washington D.C., 10 20559, where it is registered under U.S. Copyright Registration No. TXU510087. The ALIGN-2 program is publicly available through Genentech, Inc., South San Francisco, California or may be compiled from the source code provided in Table 10 below. The ALIGN-2 program should be compiled for use on a UNIX operating system, preferably digital UNIX V4.0D. All sequence comparison parameters are set by the ALIGN-2 program and do not vary.

15 In situations where ALIGN-2 is employed for amino acid sequence comparisons, the % amino acid sequence identity of a given amino acid sequence A to, with, or against a given amino acid sequence B (which can alternatively be phrased as a given amino acid sequence A that has or comprises a certain % amino acid sequence identity to, with, or against a given amino acid sequence B) is calculated as follows:

20 $100 \text{ times the fraction } X/Y$

where X is the number of amino acid residues scored as identical matches by the sequence alignment program ALIGN-2 in that program's alignment of A and B, and where Y is the total number of amino acid residues in B. It will be appreciated that where the length of amino acid sequence A is not equal to the length of amino acid 25 sequence B, the % amino acid sequence identity of A to B will not equal the % amino acid sequence identity of B to A. As examples of % amino acid sequence identity calculations using this method, Tables 11 and 12 below demonstrate how to calculate the % amino acid sequence identity of the amino acid sequence designated "Comparison Protein" to the amino acid sequence designated "PRO".

30 Unless specifically stated otherwise, all % amino acid sequence identity values used herein are obtained as described in the immediately preceding paragraph using the ALIGN-2 computer program. However, % amino acid sequence identity values may also be obtained as described below by using the WU-BLAST-2 computer program (Altschul et al., *Methods in Enzymology* 266:460-480 (1996)). Most of the WU-BLAST-2 search parameters are set to the default values. Those not set to default values, i.e., the adjustable parameters, are set with the following values: overlap span = 1, overlap fraction = 0.125, word threshold (T) = 11, and scoring 35 matrix = BLOSUM62. When WU-BLAST-2 is employed, a % amino acid sequence identity value is determined by dividing (a) the number of matching identical amino acid residues between the amino acid sequence of the PRO polypeptide of interest having a sequence derived from the native PRO polypeptide and the

comparison amino acid sequence of interest (i.e., the sequence against which the PRO polypeptide of interest is being compared which may be a PRO variant polypeptide) as determined by WU-BLAST-2 by (b) the total number of amino acid residues of the PRO polypeptide of interest. For example, in the statement "a polypeptide comprising an the amino acid sequence A which has or having at least 80% amino acid sequence identity to the amino acid sequence B", the amino acid sequence A is the comparison amino acid sequence of interest and the 5 amino acid sequence B is the amino acid sequence of the PRO polypeptide of interest.

Percent amino acid sequence identity may also be determined using the sequence comparison program NCBI-BLAST2 (Altschul et al., Nucleic Acids Res. 25:3389-3402 (1997)). The NCBI-BLAST2 sequence comparison program may be downloaded from <http://www.ncbi.nlm.nih.gov>. NCBI-BLAST2 uses several search parameters, wherein all of those search parameters are set to default values including, for example, 10 unmask = yes, strand = all, expected occurrences = 10, minimum low complexity length = 15/5, multi-pass e-value = 0.01, constant for multi-pass = 25, dropoff for final gapped alignment = 25 and scoring matrix = BLOSUM62.

In situations where NCBI-BLAST2 is employed for amino acid sequence comparisons, the % amino acid sequence identity of a given amino acid sequence A to, with, or against a given amino acid sequence B 15 (which can alternatively be phrased as a given amino acid sequence A that has or comprises a certain % amino acid sequence identity to, with, or against a given amino acid sequence B) is calculated as follows:

$$100 \text{ times the fraction } X/Y$$

20 where X is the number of amino acid residues scored as identical matches by the sequence alignment program NCBI-BLAST2 in that program's alignment of A and B, and where Y is the total number of amino acid residues in B. It will be appreciated that where the length of amino acid sequence A is not equal to the length of amino acid sequence B, the % amino acid sequence identity of A to B will not equal the % amino acid sequence identity of B to A.

25 "PRO variant polynucleotide" or "PRO variant nucleic acid sequence" means a nucleic acid molecule which encodes an active PRO polypeptide as defined below and which has at least about 80% nucleic acid sequence identity with a nucleotide acid sequence encoding a full-length native sequence PRO polypeptide sequence as disclosed herein, a full-length native sequence PRO polypeptide sequence lacking the signal peptide as disclosed herein, an extracellular domain of a PRO polypeptide, with or without the signal peptide, as 30 disclosed herein or any other fragment of a full-length PRO polypeptide sequence as disclosed herein. Ordinarily, a PRO variant polynucleotide will have at least about 80% nucleic acid sequence identity, more preferably at least about 81% nucleic acid sequence identity, more preferably at least about 82% nucleic acid sequence identity, more preferably at least about 83% nucleic acid sequence identity, more preferably at least about 84% nucleic acid sequence identity, more preferably at least about 85% nucleic acid sequence identity, 35 more preferably at least about 86% nucleic acid sequence identity, more preferably at least about 87% nucleic acid sequence identity, more preferably at least about 88% nucleic acid sequence identity, more preferably at least about 89% nucleic acid sequence identity, more preferably at least about 90% nucleic acid sequence

identity, more preferably at least about 91% nucleic acid sequence identity, more preferably at least about 92% nucleic acid sequence identity, more preferably at least about 93% nucleic acid sequence identity, more preferably at least about 94% nucleic acid sequence identity, more preferably at least about 95% nucleic acid sequence identity, more preferably at least about 96% nucleic acid sequence identity, more preferably at least about 97% nucleic acid sequence identity, more preferably at least about 98% nucleic acid sequence identity and 5 yet more preferably at least about 99% nucleic acid sequence identity with the nucleic acid sequence encoding a full-length native sequence PRO polypeptide sequence as disclosed herein, a full-length native sequence PRO polypeptide sequence lacking the signal peptide as disclosed herein, an extracellular domain of a PRO polypeptide, with or without the signal sequence, as disclosed herein or any other fragment of a full-length PRO polypeptide sequence as disclosed herein. Variants do not encompass the native nucleotide sequence.

10 Ordinarily, PRO variant polynucleotides are at least about 30 nucleotides in length, often at least about 60 nucleotides in length, more often at least about 90 nucleotides in length, more often at least about 120 nucleotides in length, more often at least about 150 nucleotides in length, more often at least about 180 nucleotides in length, more often at least about 210 nucleotides in length, more often at least about 240 nucleotides in length, more often at least about 270 nucleotides in length, more often at least about 300 15 nucleotides in length, more often at least about 450 nucleotides in length, more often at least about 600 nucleotides in length, more often at least about 900 nucleotides in length, or more.

20 "Percent (%) nucleic acid sequence identity" with respect to PRO-encoding nucleic acid sequences identified herein is defined as the percentage of nucleotides in a candidate sequence that are identical with the nucleotides in the PRO nucleic acid sequence of interest, after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent sequence identity. Alignment for purposes of determining percent nucleic acid sequence identity can be achieved in various ways that are within the skill in the art, for instance, using publicly available computer software such as BLAST, BLAST-2, ALIGN or Megalign (DNASTAR) software. For purposes herein, however, % nucleic acid sequence identity values are generated using the sequence comparison computer program ALIGN-2, wherein the complete source code for the ALIGN-2 program 25 is provided in Table 10 below. The ALIGN-2 sequence comparison computer program was authored by Genentech, Inc. and the source code shown in Table 10 below has been filed with user documentation in the U.S. Copyright Office, Washington D.C., 20559, where it is registered under U.S. Copyright Registration No. TXU510087. The ALIGN-2 program is publicly available through Genentech, Inc., South San Francisco, California or may be compiled from the source code provided in Table 10 below. The ALIGN-2 program should 30 be compiled for use on a UNIX operating system, preferably digital UNIX V4.0D. All sequence comparison parameters are set by the ALIGN-2 program and do not vary.

35 In situations where ALIGN-2 is employed for nucleic acid sequence comparisons, the % nucleic acid sequence identity of a given nucleic acid sequence C to, with, or against a given nucleic acid sequence D (which can alternatively be phrased as a given nucleic acid sequence C that has or comprises a certain % nucleic acid sequence identity to, with, or against a given nucleic acid sequence D) is calculated as follows:

100 times the fraction W/Z

where W is the number of nucleotides scored as identical matches by the sequence alignment program ALIGN-2 in that program's alignment of C and D, and where Z is the total number of nucleotides in D. It will be appreciated that where the length of nucleic acid sequence C is not equal to the length of nucleic acid sequence D, the % nucleic acid sequence identity of C to D will not equal the % nucleic acid sequence identity of D to C. As examples of % nucleic acid sequence identity calculations, Tables 13 and 14 below demonstrate how to 5 calculate the % nucleic acid sequence identity of the nucleic acid sequence designated "Comparison DNA" to the nucleic acid sequence designated "PRO-DNA".

Unless specifically stated otherwise, all % nucleic acid sequence identity values used herein are obtained as described in the immediately preceding paragraph using the ALIGN-2 computer program. However, % nucleic acid sequence identity values may also be obtained as described below by using the WU-BLAST-2 10 computer program (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Most of the WU-BLAST-2 search parameters are set to the default values. Those not set to default values, i.e., the adjustable parameters, are set with the following values: overlap span = 1, overlap fraction = 0.125, word threshold (T) = 11, and scoring matrix = BLOSUM62. When WU-BLAST-2 is employed, a % nucleic acid sequence identity value is determined by dividing (a) the number of matching identical nucleotides between the nucleic acid sequence of 15 the PRO polypeptide-encoding nucleic acid molecule of interest having a sequence derived from the native sequence PRO polypeptide-encoding nucleic acid and the comparison nucleic acid molecule of interest (i.e., the sequence against which the PRO polypeptide-encoding nucleic acid molecule of interest is being compared which may be a variant PRO polynucleotide) as determined by WU-BLAST-2 by (b) the total number of nucleotides of the PRO polypeptide-encoding nucleic acid molecule of interest. For example, in the statement "an isolated 20 nucleic acid molecule comprising a nucleic acid sequence A which has or having at least 80% nucleic acid sequence identity to the nucleic acid sequence B", the nucleic acid sequence A is the comparison nucleic acid molecule of interest and the nucleic acid sequence B is the nucleic acid sequence of the PRO polypeptide-encoding nucleic acid molecule of interest.

Percent nucleic acid sequence identity may also be determined using the sequence comparison program 25 NCBI-BLAST2 (Altschul et al., Nucleic Acids Res. 25:3389-3402 (1997)). The NCBI-BLAST2 sequence comparison program may be downloaded from <http://www.ncbi.nlm.nih.gov>. NCBI-BLAST2 uses several search parameters, wherein all of those search parameters are set to default values including, for example, unmask = yes, strand = all, expected occurrences = 10, minimum low complexity length = 15/5, multi-pass e-value = 0.01, constant for multi-pass = 25, dropoff for final gapped alignment = 25 and scoring matrix = 30 BLOSUM62.

In situations where NCBI-BLAST2 is employed for sequence comparisons, the % nucleic acid sequence identity of a given nucleic acid sequence C to, with, or against a given nucleic acid sequence D (which can alternatively be phrased as a given nucleic acid sequence C that has or comprises a certain % nucleic acid sequence identity to, with, or against a given nucleic acid sequence D) is calculated as follows:

35

100 times the fraction W/Z

where W is the number of nucleotides scored as identical matches by the sequence alignment program NCBI-BLAST2 in that program's alignment of C and D, and where Z is the total number of nucleotides in D. It will be appreciated that where the length of nucleic acid sequence C is not equal to the length of nucleic acid sequence D, the % nucleic acid sequence identity of C to D will not equal the % nucleic acid sequence identity of D to C.

5 In other embodiments, PRO variant polynucleotides are nucleic acid molecules that encode an active PRO polypeptide and which are capable of hybridizing, preferably under stringent hybridization and wash conditions, to nucleotide sequences encoding a full-length PRO polypeptide as disclosed herein. PRO variant polypeptides may be those that are encoded by a PRO variant polynucleotide.

The term "positives", in the context of sequence comparison performed as described above, includes 10 residues in the sequences compared that are not identical but have similar properties (e.g. as a result of conservative substitutions, see Table 1 below). For purposes herein, the % value of positives is determined by dividing (a) the number of amino acid residues scoring a positive value between the PRO polypeptide amino acid sequence of interest having a sequence derived from the native PRO polypeptide sequence and the comparison amino acid sequence of interest (i.e., the amino acid sequence against which the PRO polypeptide sequence is 15 being compared) as determined in the BLOSUM62 matrix of WU-BLAST-2 by (b) the total number of amino acid residues of the PRO polypeptide of interest.

Unless specifically stated otherwise, the % value of positives is calculated as described in the immediately preceding paragraph. However, in the context of the amino acid sequence identity comparisons performed as described for ALIGN-2 and NCBI-BLAST2 above, includes amino acid residues in the sequences 20 compared that are not only identical, but also those that have similar properties. Amino acid residues that score a positive value to an amino acid residue of interest are those that are either identical to the amino acid residue of interest or are a preferred substitution (as defined in Table 1 below) of the amino acid residue of interest.

For amino acid sequence comparisons using ALIGN-2 or NCBI-BLAST2, the % value of positives of a given amino acid sequence A to, with, or against a given amino acid sequence B (which can alternatively be 25 phrased as a given amino acid sequence A that has or comprises a certain % positives to, with, or against a given amino acid sequence B) is calculated as follows:

$$100 \text{ times the fraction } X/Y$$

30 where X is the number of amino acid residues scoring a positive value as defined above by the sequence alignment program ALIGN-2 or NCBI-BLAST2 in that program's alignment of A and B, and where Y is the total number of amino acid residues in B. It will be appreciated that where the length of amino acid sequence A is not equal to the length of amino acid sequence B, the % positives of A to B will not equal the % positives of B to A.

35 "Isolated," when used to describe the various polypeptides disclosed herein, means polypeptide that has been identified and separated and/or recovered from a component of its natural environment. Contaminant components of its natural environment are materials that would typically interfere with diagnostic or therapeutic

uses for the polypeptide, and may include enzymes, hormones, and other proteinaceous or non-proteinaceous solutes. In preferred embodiments, the polypeptide will be purified (1) to a degree sufficient to obtain at least 15 residues of N-terminal or internal amino acid sequence by use of a spinning cup sequenator, or (2) to homogeneity by SDS-PAGE under non-reducing or reducing conditions using Coomassie blue or, preferably, silver stain. Isolated polypeptide includes polypeptide *in situ* within recombinant cells, since at least one 5 component of the PRO polypeptide natural environment will not be present. Ordinarily, however, isolated polypeptide will be prepared by at least one purification step.

An "isolated" PRO polypeptide-encoding nucleic acid is a nucleic acid molecule that is identified and separated from at least one contaminant nucleic acid molecule with which it is ordinarily associated in the natural source of the PRO polypeptide nucleic acid. An isolated PRO polypeptide nucleic acid molecule is other than 10 in the form or setting in which it is found in nature. Isolated PRO polypeptide nucleic acid molecules therefore are distinguished from the specific PRO polypeptide nucleic acid molecule as it exists in natural cells. However, an isolated PRO polypeptide nucleic acid molecule includes PRO polypeptide nucleic acid molecules contained in cells that ordinarily express the PRO polypeptide where, for example, the nucleic acid molecule is in a chromosomal location different from that of natural cells.

15 The term "control sequences" refers to DNA sequences necessary for the expression of an operably linked coding sequence in a particular host organism. The control sequences that are suitable for prokaryotes, for example, include a promoter, optionally an operator sequence, and a ribosome binding site. Eukaryotic cells are known to utilize promoters, polyadenylation signals, and enhancers.

20 Nucleic acid is "operably linked" when it is placed into a functional relationship with another nucleic acid sequence. For example, DNA for a presequence or secretory leader is operably linked to DNA for a polypeptide if it is expressed as a preprotein that participates in the secretion of the polypeptide; a promoter or enhancer is operably linked to a coding sequence if it affects the transcription of the sequence; or a ribosome binding site is operably linked to a coding sequence if it is positioned so as to facilitate translation. Generally, 25 "operably linked" means that the DNA sequences being linked are contiguous, and, in the case of a secretory leader, contiguous and in reading phase. However, enhancers do not have to be contiguous. Linking is accomplished by ligation at convenient restriction sites. If such sites do not exist, the synthetic oligonucleotide adaptors or linkers are used in accordance with conventional practice.

30 The term "antibody" is used in the broadest sense and specifically covers, for example, single anti-PRO monoclonal antibodies (including agonist, antagonist, and neutralizing antibodies), anti-PRO antibody compositions with polyepitopic specificity, single chain anti-PRO antibodies, and fragments of anti-PRO antibodies (see below). The term "monoclonal antibody" as used herein refers to an antibody obtained from a population of substantially homogeneous antibodies, i.e., the individual antibodies comprising the population are identical except for possible naturally-occurring mutations that may be present in minor amounts.

35 "Stringency" of hybridization reactions is readily determinable by one of ordinary skill in the art, and generally is an empirical calculation dependent upon probe length, washing temperature, and salt concentration. In general, longer probes require higher temperatures for proper annealing, while shorter probes need lower temperatures. Hybridization generally depends on the ability of denatured DNA to reanneal when

complementary strands are present in an environment below their melting temperature. The higher the degree of desired homology between the probe and hybridizable sequence, the higher the relative temperature which can be used. As a result, it follows that higher relative temperatures would tend to make the reaction conditions more stringent, while lower temperatures less so. For additional details and explanation of stringency of hybridization reactions, see Ausubel et al., Current Protocols in Molecular Biology, Wiley Interscience Publishers, (1995).

"Stringent conditions" or "high stringency conditions", as defined herein, may be identified by those that: (1) employ low ionic strength and high temperature for washing, for example 0.015 M sodium chloride/0.0015 M sodium citrate/0.1% sodium dodecyl sulfate at 50°C; (2) employ during hybridization a denaturing agent, such as formamide, for example, 50% (v/v) formamide with 0.1% bovine serum albumin/0.1% Ficoll/0.1% polyvinylpyrrolidone/50mM sodium phosphate buffer at pH 6.5 with 750 mM sodium chloride, 75 mM sodium citrate at 42°C; or (3) employ 50% formamide, 5 x SSC (0.75 M NaCl, 0.075 M sodium citrate), 50 mM sodium phosphate (pH 6.8), 0.1% sodium pyrophosphate, 5 x Denhardt's solution, sonicated salmon sperm DNA (50 µg/ml), 0.1% SDS, and 10% dextran sulfate at 42°C, with washes at 42°C in 0.2 x SSC (sodium chloride/sodium citrate) and 50% formamide at 55°C, followed by a high-stringency wash 15 consisting of 0.1 x SSC containing EDTA at 55°C.

"Moderately stringent conditions" may be identified as described by Sambrook et al., Molecular Cloning: A Laboratory Manual, New York: Cold Spring Harbor Press, 1989, and include the use of washing solution and hybridization conditions (e.g., temperature, ionic strength and %SDS) less stringent than those described above. An example of moderately stringent conditions is overnight incubation at 37°C in a solution 20 comprising: 20% formamide, 5 x SSC (150 mM NaCl, 15 mM trisodium citrate), 50 mM sodium phosphate (pH 7.6), 5 x Denhardt's solution, 10% dextran sulfate, and 20 mg/ml denatured sheared salmon sperm DNA, followed by washing the filters in 1 x SSC at about 37-50°C. The skilled artisan will recognize how to adjust the temperature, ionic strength, etc. as necessary to accommodate factors such as probe length and the like.

The term "epitope tagged" when used herein refers to a chimeric polypeptide comprising a PRO 25 polypeptide fused to a "tag polypeptide". The tag polypeptide has enough residues to provide an epitope against which an antibody can be made, yet is short enough such that it does not interfere with activity of the polypeptide to which it is fused. The tag polypeptide preferably also is fairly unique so that the antibody does not substantially cross-react with other epitopes. Suitable tag polypeptides generally have at least six amino acid residues and usually between about 8 and 50 amino acid residues (preferably, between about 10 and 20 amino 30 acid residues).

As used herein, the term "immunoadhesin" designates antibody-like molecules which combine the binding specificity of a heterologous protein (an "adhesin") with the effector functions of immunoglobulin constant domains. Structurally, the immunoadhesins comprise a fusion of an amino acid sequence with the desired binding specificity which is other than the antigen recognition and binding site of an antibody (i.e., is 35 "heterologous"), and an immunoglobulin constant domain sequence. The adhesin part of an immunoadhesin molecule typically is a contiguous amino acid sequence comprising at least the binding site of a receptor or a ligand. The immunoglobulin constant domain sequence in the immunoadhesin may be obtained from any

immunoglobulin, such as IgG-1, IgG-2, IgG-3, or IgG-4 subtypes, IgA (including IgA-1 and IgA-2); IgE, IgD or IgM.

"Active" or "activity" for the purposes herein refers to form(s) of a PRO polypeptide which retain a biological and/or an immunological activity of native or naturally-occurring PRO, wherein "biological" activity refers to a biological function (either inhibitory or stimulatory) caused by a native or naturally-occurring PRO other than the ability to induce the production of an antibody against an antigenic epitope possessed by a native or naturally-occurring PRO and an "immunological" activity refers to the ability to induce the production of an antibody against an antigenic epitope possessed by a native or naturally-occurring PRO.

The term "antagonist" is used in the broadest sense, and includes any molecule that partially or fully blocks, inhibits, or neutralizes a biological activity of a native PRO polypeptide disclosed herein. In a similar manner, the term "agonist" is used in the broadest sense and includes any molecule that mimics a biological activity of a native PRO polypeptide disclosed herein. Suitable agonist or antagonist molecules specifically include agonist or antagonist antibodies or antibody fragments, fragments or amino acid sequence variants of native PRO polypeptides, peptides, antisense oligonucleotides, small organic molecules, etc. Methods for identifying agonists or antagonists of a PRO polypeptide may comprise contacting a PRO polypeptide with a candidate agonist or antagonist molecule and measuring a detectable change in one or more biological activities normally associated with the PRO polypeptide.

"Treatment" refers to both therapeutic treatment and prophylactic or preventative measures, wherein the object is to prevent or slow down (lessen) the targeted pathologic condition or disorder. Those in need of treatment include those already with the disorder as well as those prone to have the disorder or those in whom the disorder is to be prevented.

"Chronic" administration refers to administration of the agent(s) in a continuous mode as opposed to an acute mode, so as to maintain the initial therapeutic effect (activity) for an extended period of time. "Intermittent" administration is treatment that is not consecutively done without interruption, but rather is cyclic in nature.

"Mammal" for purposes of treatment refers to any animal classified as a mammal, including humans, domestic and farm animals, and zoo, sports, or pet animals, such as dogs, cats, cattle, horses, sheep, pigs, goats, rabbits, etc. Preferably, the mammal is human.

Administration "in combination with" one or more further therapeutic agents includes simultaneous (concurrent) and consecutive administration in any order.

"Carriers" as used herein include pharmaceutically acceptable carriers, excipients, or stabilizers which are nontoxic to the cell or mammal being exposed thereto at the dosages and concentrations employed. Often the physiologically acceptable carrier is an aqueous pH buffered solution. Examples of physiologically acceptable carriers include buffers such as phosphate, citrate, and other organic acids; antioxidants including ascorbic acid; low molecular weight (less than about 10 residues) polypeptide; proteins, such as serum albumin, gelatin, or immunoglobulins; hydrophilic polymers such as polyvinylpyrrolidone; amino acids such as glycine, glutamine, asparagine, arginine or lysine; monosaccharides, disaccharides, and other carbohydrates including glucose, mannose, or dextrans; chelating agents such as EDTA; sugar alcohols such as mannitol or sorbitol; salt-

forming counterions such as sodium; and/or nonionic surfactants such as TWEEN™, polyethylene glycol (PEG), and PLURONICS™.

"Antibody fragments" comprise a portion of an intact antibody, preferably the antigen binding or variable region of the intact antibody. Examples of antibody fragments include Fab, Fab', F(ab')₂, and Fv fragments; diabodies; linear antibodies (Zapata et al., Protein Eng. 8(10): 1057-1062 [1995]); single-chain antibody molecules; and multispecific antibodies formed from antibody fragments.

Papain digestion of antibodies produces two identical antigen-binding fragments, called "Fab" fragments, each with a single antigen-binding site, and a residual "Fc" fragment, a designation reflecting the ability to crystallize readily. Pepsin treatment yields an F(ab')₂ fragment that has two antigen-combining sites and is still capable of cross-linking antigen.

"Fv" is the minimum antibody fragment which contains a complete antigen-recognition and -binding site. This region consists of a dimer of one heavy- and one light-chain variable domain in tight, non-covalent association. It is in this configuration that the three CDRs of each variable domain interact to define an antigen-binding site on the surface of the VH-VL dimer. Collectively, the six CDRs confer antigen-binding specificity to the antibody. However, even a single variable domain (or half of an Fv comprising only three CDRs specific for an antigen) has the ability to recognize and bind antigen, although at a lower affinity than the entire binding site.

The Fab fragment also contains the constant domain of the light chain and the first constant domain (CH1) of the heavy chain. Fab fragments differ from Fab' fragments by the addition of a few residues at the carboxy terminus of the heavy chain CH1 domain including one or more cysteines from the antibody hinge region. Fab'-SH is the designation herein for Fab' in which the cysteine residue(s) of the constant domains bear a free thiol group. F(ab')₂ antibody fragments originally were produced as pairs of Fab' fragments which have hinge cysteines between them. Other chemical couplings of antibody fragments are also known.

The "light chains" of antibodies (immunoglobulins) from any vertebrate species can be assigned to one of two clearly distinct types, called kappa and lambda, based on the amino acid sequences of their constant domains.

Depending on the amino acid sequence of the constant domain of their heavy chains, immunoglobulins can be assigned to different classes. There are five major classes of immunoglobulins: IgA, IgD, IgE, IgG, and IgM, and several of these may be further divided into subclasses (isotypes), e.g., IgG1, IgG2, IgG3, IgG4, IgA, and IgA2.

"Single-chain Fv" or "sFv" antibody fragments comprise the VH and VL domains of antibody, wherein these domains are present in a single polypeptide chain. Preferably, the Fv polypeptide further comprises a polypeptide linker between the VH and VL domains which enables the sFv to form the desired structure for antigen binding. For a review of sFv, see Pluckthun in The Pharmacology of Monoclonal Antibodies, vol. 113, Rosenberg and Moore eds., Springer-Verlag, New York, pp. 269-315 (1994).

The term "diabodies" refers to small antibody fragments with two antigen-binding sites, which fragments comprise a heavy-chain variable domain (VH) connected to a light-chain variable domain (VL) in the same polypeptide chain (VH - VL). By using a linker that is too short to allow pairing between the two domains

on the same chain, the domains are forced to pair with the complementary domains of another chain and create two antigen-binding sites. Diabodies are described more fully in, for example, EP 404,097; WO 93/11161; and Hollinger et al., Proc. Natl. Acad. Sci. USA, 90:6444-6448 (1993).

An "isolated" antibody is one which has been identified and separated and/or recovered from a component of its natural environment. Contaminant components of its natural environment are materials which 5 would interfere with diagnostic or therapeutic uses for the antibody, and may include enzymes, hormones, and other proteinaceous or nonproteinaceous solutes. In preferred embodiments, the antibody will be purified (1) to greater than 95% by weight of antibody as determined by the Lowry method, and most preferably more than 99% by weight, (2) to a degree sufficient to obtain at least 15 residues of N-terminal or internal amino acid sequence by use of a spinning cup sequenator, or (3) to homogeneity by SDS-PAGE under reducing or 10 nonreducing conditions using Coomassie blue or, preferably, silver stain. Isolated antibody includes the antibody in situ within recombinant cells since at least one component of the antibody's natural environment will not be present. Ordinarily, however, isolated antibody will be prepared by at least one purification step.

The word "label" when used herein refers to a detectable compound or composition which is conjugated directly or indirectly to the antibody so as to generate a "labeled" antibody. The label may be detectable by itself 15 (e.g. radioisotope labels or fluorescent labels) or, in the case of an enzymatic label, may catalyze chemical alteration of a substrate compound or composition which is detectable.

By "solid phase" is meant a non-aqueous matrix to which the antibody of the present invention can adhere. Examples of solid phases encompassed herein include those formed partially or entirely of glass (e.g., controlled pore glass), polysaccharides (e.g., agarose), polyacrylamides, polystyrene, polyvinyl alcohol and 20 silicones. In certain embodiments, depending on the context, the solid phase can comprise the well of an assay plate; in others it is a purification column (e.g., an affinity chromatography column). This term also includes a discontinuous solid phase of discrete particles, such as those described in U.S. Patent No. 4,275,149.

A "liposome" is a small vesicle composed of various types of lipids, phospholipids and/or surfactant which is useful for delivery of a drug (such as a PRO polypeptide or antibody thereto) to a mammal. The 25 components of the liposome are commonly arranged in a bilayer formation, similar to the lipid arrangement of biological membranes.

A "small molecule" is defined herein to have a molecular weight below about 500 Daltons.

"PRO317-associated disorder" refers to a pathological condition or disease wherein PRO317 is over- or underexpressed. Such disorders include diseases of the female genital tract or of the endometrium of a 30 mammal, including hyperplasia, endometritis, endometriosis, wherein the patient is at risk for infertility due to endometrial factor, endometrioma, and endometrial cancer, especially those diseases involving abnormal bleeding such as a gynecological disease. They also include diseases involving angiogenesis, wherein the angiogenesis results in a pathological condition, such as cancer involving solid tumors (the therapy for the disorder would result in decreased vascularization and a decline in growth and metastasis of a variety of tumors). Alternatively, 35 the angiogenesis may be beneficial, such as for ischemia, especially coronary ischemia. Hence, these disorders include those found in patients whose hearts are functioning but who have a blocked blood supply due to atherosclerotic coronary artery disease, and those with a functioning but underperfused heart, including patients

with coronary arterial disease who are not optimal candidates for angioplasty and coronary artery by-pass surgery. The disorders also include diseases involving the kidney or originating from the kidney tissue, such as polycystic kidney disease and chronic and acute renal failure.

II. Compositions and Methods of the Invention

5 A. Full-Length PRO Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO polypeptides. In particular, cDNAs encoding various PRO polypeptides have been identified and isolated, as disclosed in further detail in the Examples below. It is noted that proteins produced in separate expression rounds may be given different PRO numbers but the UNQ number 10 is unique for any given DNA and the encoded protein, and will not be changed. However, for sake of simplicity, in the present specification the protein encoded by the full length native nucleic acid molecules disclosed herein as well as all further native homologues and variants included in the foregoing definition of PRO, will be referred to as "PRO/number", regardless of their origin or mode of preparation.

As disclosed in the Examples below, various cDNA clones have been deposited with the ATCC. The 15 actual nucleotide sequences of those clones can readily be determined by the skilled artisan by sequencing of the deposited clone using routine methods in the art. The predicted amino acid sequence can be determined from the nucleotide sequence using routine skill. For the PRO polypeptides and encoding nucleic acids described herein, Applicants have identified what is believed to be the reading frame best identifiable with the sequence information available at the time.

20

1. Full-length PRO211 and PRO217 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO211 and PRO217. In particular, Applicants have identified and isolated cDNA encoding PRO211 and PRO217 polypeptides, as disclosed in further detail in the Examples 25 below. Using BLAST (FastA format) sequence alignment computer programs, Applicants found that cDNA sequences encoding full-length native sequence PRO211 and PRO217 have homologies to known proteins having EGF-like domains. Specifically, the cDNA sequence DNA32292-1131 (Figure 1, SEQ ID NO:1) has certain identify and a Blast score of 209 with PAC6_RAT and certain identify and a Blast score of 206 with Fibulin-1, isoform c precursor. The cDNA sequence DNA33094-1131 (Figure 3, SEQ ID NO:3) has 36% identity and 30 a Blast score of 336 with eastern newt tenascin, and 37% identity and a Blast score of 331 with human tenascin-X precursor. Accordingly, it is presently believed that PRO211 and PRO217 polypeptides disclosed in the present application are newly identified members of the EGF-like family and possesses properties typical of the EGF-like protein family.

35 2. Full-length PRO230 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO230. In particular, Applicants have identified and isolated cDNA

encoding a PRO230 polypeptide, as disclosed in further detail in the Examples below. Using known programs such as BLAST and FastA sequence alignment computer programs, Applicants found that a cDNA sequence encoding full-length native sequence PRO230 has 48% amino acid identity with the rabbit tubulointerstitial nephritis antigen precursor. Accordingly, it is presently believed that PRO230 polypeptide disclosed in the present application is a newly identified member of the tubulointerstitial nephritis antigen family and possesses 5 the ability to be recognized by human autoantibodies in certain forms of tubulointerstitial nephritis.

3. Full-length PRO232 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO232. In particular, Applicants have identified and isolated cDNA 10 encoding a PRO232 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that a portion of the full-length native sequence PRO232 (shown in Figure 9 and SEQ ID NO:18) has 35% sequence identity with a stem cell surface antigen from Gallus gallus. Accordingly, it is presently believed that the PRO232 polypeptide disclosed in the present application may be a newly identified stem cell antigen.

15

4. Full-length PRO187 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO187. In particular, Applicants have identified and isolated cDNA 20 encoding a PRO187 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that a full-length native sequence PRO187 (shown in Figure 15) has 74% amino acid sequence identity and BLAST score of 310 with various androgen-induced growth factors and FGF-8. Accordingly, it is presently believed that PRO187 polypeptide disclosed in the present application is a newly identified member of the FGF-8 protein family and may possess identify activity or property typical of the FGF-8-like protein family.

25

5. Full-length PRO265 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO265. In particular, Applicants have identified and isolated cDNA 30 encoding a PRO265 polypeptide, as disclosed in further detail in the Examples below. Using programs such as BLAST and FastA sequence alignment computer programs, Applicants found that various portions of the PRO265 polypeptide have significant homology with the fibromodulin protein and fibromodulin precursor protein. Applicants have also found that the DNA encoding the PRO265 polypeptide has significant homology with platelet glycoprotein V, a member of the leucine rich related protein family involved in skin and wound repair. Accordingly, it is presently believed that PRO265 polypeptide disclosed in the present application is a 35 newly identified member of the leucine rich repeat family and possesses protein protein binding capabilities, as well as be involved in skin and wound repair as typical of this family.

6. Full-length PRO219 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO219. In particular, Applicants have identified and isolated cDNA encoding a PRO219 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that various portions of the PRO219 polypeptide have 5 significant homology with the mouse and human matrilin-2 precursor polypeptides. Accordingly, it is presently believed that PRO219 polypeptide disclosed in the present application is related to the matrilin-2 precursor polypeptide.

7. Full-length PRO246 Polypeptides

10 The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO246. In particular, Applicants have identified and isolated cDNA encoding a PRO246 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that a portion of the PRO246 polypeptide has significant homology with the human cell surface protein HCAR. Accordingly, it is presently believed that 15 PRO246 polypeptide disclosed in the present application may be a newly identified membrane-bound virus receptor or tumor cell-specific antigen.

8. Full-length PRO228 Polypeptides

20 The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO228. In particular, Applicants have identified and isolated cDNA encoding a PRO228 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that various portions of the PRO228 polypeptide have significant homology with the EMR1 protein. Applicants have also found that the DNA encoding the PRO228 25 polypeptide has significant homology with latrophilin, macrophage-restricted cell surface glycoprotein, B0457.1 and leucocyte antigen CD97 precursor. Accordingly, it is presently believed that PRO228 polypeptide disclosed in the present application is a newly identified member of the seven transmembrane superfamily and possesses characteristics and functional properties typical of this family. In particular, it is believed that PRO228 is a new member of the subgroup within this family to which CD97 and EMR1 belong.

30 9. Full-length PRO533 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO533. In particular, Applicants have identified and isolated cDNA encoding a PRO533 polypeptide, as disclosed in further detail in the Examples below. Using BLAST-2 and FastA sequence alignment computer programs, Applicants found that a full-length native sequence PRO533 35 (shown in Figure 22 and SEQ ID NO:59) has a Blast score of 509 and 53% amino acid sequence identity with fibroblast growth factor (FGF). Accordingly, it is presently believed that PRO533 disclosed in the present application is a newly identified member of the fibroblast growth factor family and may possess activity typical

of such polypeptides.

10. **Full-length PRO245 Polypeptides**

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO245. In particular, Applicants have identified and isolated cDNA 5 encoding a PRO245 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that a portion of the amino acid sequence of the PRO245 polypeptide has 60% amino acid identity with the human c-myb protein. Accordingly, it is presently believed that the PRO245 polypeptide disclosed in the present application may be a newly identified member of the transmembrane protein tyrosine kinase family.

10

11. **Full-length PRO220, PRO221 and PRO227 Polypeptides**

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO220, PRO221 and PRO227. In particular, Applicants have identified and isolated cDNAs encoding a PRO220, PRO221 and PRO227 polypeptide, respectively, as disclosed in further 15 detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, PRO220 has amino acid identity with the amino acid sequence of a leucine rich protein wherein the identity is 87%. PRO220 additionally has amino acid identity with the neuronal leucine rich protein wherein the identity is 55%. The neuronal leucine rich protein is further described in Taguchi, *et al.*, *Mol. Brain Res.*, 35:31-40 (1996).

20 PRO221 has amino acid identity with the SLIT protein precursor, wherein different portions of these two proteins have the respective percent identities of 39%, 38%, 34%, 31%, and 30%.

PRO227 has amino acid identity with the amino acid sequence of platelet glycoprotein V precursor. The same results were obtained for human glycoprotein V. Different portions of these two proteins show the following percent identities of 30%, 28%, 28%, 31%, 35%, 39% and 27%.

25 Accordingly, it is presently believed that PRO220, PRO221 and PRO227 polypeptides disclosed in the present application are newly identified members of the leucine rich repeat protein superfamily and that each possesses protein-protein binding capabilities typical of the leucine rich repeat protein superfamily. It is also believed that they have capabilities similar to those of SLIT, the leucine rich repeat protein and human glycoprotein V.

30

12. **Full-length PRO258 Polypeptides**

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO258. In particular, Applicants have identified and isolated cDNA 35 encoding a PRO258 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that various portions of the PRO258 polypeptide have significant homology with the CRTAM and poliovirus receptors. Accordingly, it is presently believed that PRO258 polypeptide disclosed in the present application is a newly identified member of the Ig superfamily and possesses virus receptor capabilities or regulates immune function as typical of this family.

13. Full-length PRO266 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO266. In particular, Applicants have identified and isolated cDNA encoding a PRO266 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that various portions of the PRO266 polypeptide have 5 significant homology with the SLIT protein from Drosophila. Accordingly, it is presently believed that PRO266 polypeptide disclosed in the present application is a newly identified member of the leucine rich repeat family and possesses ligand-ligand binding activity and neuronal development typical of this family. SLIT has been shown to be useful in the study and treatment of Alzheimer's disease, *supra*, and thus, PRO266 may have involvement in the study and cure of this disease.

10

14. Full-length PRO269 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO269. In particular, Applicants have identified and isolated cDNA encoding a PRO269 polypeptide, as disclosed in further detail in the Examples below. Using BLAST, FastA 15 and sequence alignment computer programs, Applicants found that the amino acid sequence encoded by nucleotides 314 to 1783 of the full-length native sequence PRO269 (shown in Figure 35 and SEQ ID NO:95) has significant homology to human urinary thrombomodulin and various thrombomodulin analogues respectively, to which it was aligned. Accordingly, it is presently believed that PRO269 polypeptide disclosed in the present application is a newly identified member of the thrombomodulin family.

20

15. Full-length PRO287 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO287. In particular, Applicants have identified and isolated cDNA encoding a PRO287 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA 25 sequence alignment computer programs, Applicants found that various portions of the PRO287 polypeptide have significant homology with the type 1 procollagen C-proteinase enhancer protein precursor and type 1 procollagen C-proteinase enhancer protein. Accordingly, it is presently believed that PRO287 polypeptide disclosed in the present application is a newly identified member of the C-proteinase enhancer protein family.

30

16. Full-length PRO214 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO214. In particular, Applicants have identified and isolated cDNA encoding a PRO214 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that a full-length native sequence PRO214 polypeptide 35 (shown in Figure 40 and SEQ ID NO:109) has 49% amino acid sequence identity with HT protein, a known member of the EGF-family. The comparison resulted in a BLAST score of 920, with 150 matching nucleotides. Accordingly, it is presently believed that the PRO214 polypeptide disclosed in the present application is a newly

identified member of the family comprising EGF domains and may possess activities or properties typical of the EGF-domain containing family.

17. Full-length PRO317 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO317. In particular, cDNA encoding a PRO317 polypeptide has been identified and isolated, as disclosed in further detail in the Examples below. Using BLAST™ and FastA™ sequence alignment computer programs, it was found that a full-length native-sequence PRO317 (shown in Figure 42 and SEQ ID NO:114) has 92% amino acid sequence identity with EBAF-1. Further, it is closely aligned with many other members of the TGF- superfamily.

Accordingly, it is presently believed that PRO317 disclosed in the present application is a newly identified member of the TGF- superfamily and may possess properties that are therapeutically useful in conditions of uterine bleeding, etc. Hence, PRO317 may be useful in diagnosing or treating abnormal bleeding involved in gynecological diseases, for example, to avoid or lessen the need for a hysterectomy. PRO317 may also be useful as an agent that affects angiogenesis in general, so PRO317 may be useful in anti-tumor indications, or conversely, in treating coronary ischemic conditions.

Library sources reveal that ESTs used to obtain the consensus DNA for generating PRO317 primers and probes were found in normal tissues (uterus, prostate, colon, and pancreas), in several tumors (colon, brain (twice), pancreas, and mullerian cell), and in a heart with ischemia. PRO317 has shown up in several tissues as well, but it does look to have a greater concentration in uterus. Hence, PRO317 may have a broader use by the body than EBAF-1. It is contemplated that, at least for some indications, PRO317 may have opposite effects from EBAF-1.

18. Full-length PRO301 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO301. In particular, Applicants have identified and isolated cDNA encoding a PRO301 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that a full-length native sequence PRO301 (shown in Figure 44 and SEQ ID NO:119) has a Blast score of 246 corresponding to 30% amino acid sequence identity with human A33 antigen precursor. Accordingly, it is presently believed that PRO301 disclosed in the present application is a newly identified member of the A33 antigen protein family and may be expressed in human neoplastic diseases such as colorectal cancer.

19. Full-length PRO224 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO224. In particular, Applicants have identified and isolated cDNA encoding a PRO224 polypeptide, as disclosed in further detail in the Examples below. Using known programs such as BLAST and FastA sequence alignment computer programs, Applicants found that full-length native

PRO224 (Figure 46, SEQ ID NO:127) has amino acid identity with apolipoprotein E receptor 2906 from homo sapiens. The alignments of different portions of these two polypeptides show amino acid identities of 37%, 36%, 30%, 44%, 44% and 28% respectively. Full-length native PRO224 (Figure 46, SEQ ID NO:127) also has amino acid identity with very low-density lipoprotein receptor precursor from gall. The alignments of different portions of these two polypeptides show amino acid identities of 38%, 37%, 42%, 33%, and 37% respectively.

5 Additionally, full-length native PRO224 (Figure 46, SEQ ID NO:127) has amino acid identity with the chicken oocyte receptor P95 from Gallus gallus. The alignments of different portions of these two polypeptides show amino acid identities of 38%, 37%, 42%, 33%, and 37% respectively. Moreover, full-length native PRO224 (Figure 46, SEQ ID NO:127) has amino acid identity with very low density lipoprotein receptor short form precursor from humans. The alignments of different portions of these two polypeptides show amino acid 10 identities of 32%, 38%, 34%, 45%, and 31%, respectively. Accordingly, it is presently believed that PRO224 polypeptide disclosed in the present application is a newly identified member of the low density lipoprotein receptor family and possesses the structural characteristics required to have the functional ability to recognize and endocytose low density lipoproteins typical of the low density lipoprotein receptor family. (The alignments described above used the following scoring parameters: T=7, S+65, S2=36, Matrix: BLOSUM62.)

15

20. Full-length PRO222 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO222. In particular, Applicants have identified and isolated cDNA encoding a PRO222 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA 20 sequence alignment computer programs, Applicants found that a sequence encoding full-length native sequence PRO222 (shown in Figure 48 and SEQ ID NO:132) has 25-26% amino acid identity with mouse complement factor h precursor, has 27-29% amino acid identity with complement receptor, has 25-47% amino acid identity with mouse complement C3b receptor type 2 long form precursor, has 40% amino acid identity with human hypothetical protein kiaa0247. Accordingly, it is presently believed that PRO222 polypeptide disclosed in the 25 present application is a newly identified member of the complement receptor family and possesses activity typical of the complement receptor family.

21. Full-length PRO234 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides 30 referred to in the present application as PRO234. In particular, Applicants have identified and isolated cDNA encoding a PRO234 polypeptide, as disclosed in further detail in the Examples below. Using BLAST (FastA-format) sequence alignment computer programs, Applicants found that a cDNA sequence encoding full-length native sequence PRO234 has 31% identity and Blast score of 134 with E-selectin precursor. Accordingly, it is presently believed that the PRO234 polypeptides disclosed in the present application are newly identified 35 members of the lectin/selectin family and possess activity typical of the lectin/selectin family.

22. Full-length PRO231 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO231. In particular, Applicants have identified and isolated cDNA encoding a PRO231 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that the full-length native sequence PRO231 5 polypeptide (shown in Figure 52 and SEQ ID NO:142) has 30 % and 31 % amino acid identity with human and rat prostatic acid phosphatase precursor proteins, respectively. Accordingly, it is presently believed that the PRO231 polypeptide disclosed in the present application may be a newly identified member of the acid phosphatase protein family.

10

23. Full-length PRO229 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO229. In particular, Applicants have identified and isolated cDNA encoding a PRO229 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that various portions of the PRO229 polypeptide have 15 significant homology with antigen wc1.1, M130 antigen, T cell surface glycoprotein CD6 and CD6. It also is related to Sp-alpha. Accordingly, it is presently believed that PRO229 polypeptide disclosed in the present application is a newly identified member of the family containing scavenger receptor homology, a sequence motif found in a number of proteins involved in immune function and thus possesses immune function and /or segments which resist degradation, typical of this family.

20

24. Full-length PRO238 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO238. In particular, Applicants have identified and isolated cDNA encoding a PRO238 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA 25 sequence alignment computer programs, Applicants found that various portions of the PRO238 polypeptide have significant homology with reductases, including oxidoreductase and fatty acyl-CoA reductase. Accordingly, it is presently believed that PRO238 polypeptide disclosed in the present application is a newly identified member of the reductase family and possesses reducing activity typical of the reductase family.

30

25. Full-length PRO233 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO233. In particular, Applicants have identified and isolated cDNA encoding a PRO233 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that various portions of the PRO233 polypeptide have 35 significant homology with the reductase protein. Applicants have also found that the DNA encoding the PRO233 polypeptide has significant homology with proteins from *Caenorhabditis elegans*. Accordingly, it is presently believed that PRO233 polypeptide disclosed in the present application is a newly identified member of the

reductase family and possesses the ability to effect the redox state of the cell typical of the reductase family.

26. **Full-length PRO223 Polypeptides**

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO223. In particular, Applicants have identified and isolated cDNA 5 encoding a PRO223 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that the PRO223 polypeptide has significant homology with various serine carboxypeptidase polypeptides. Accordingly, it is presently believed that PRO223 polypeptide disclosed in the present application is a newly identified serine carboxypeptidase.

10 27. **Full-length PRO235 Polypeptides**

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO235. In particular, Applicants have identified and isolated cDNA encoding a PRO235 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that various portions of the PRO235 polypeptide have 15 significant homology with the various plexin proteins. Accordingly, it is presently believed that PRO235 polypeptide disclosed in the present application is a newly identified member of the plexin family and possesses cell adhesion properties typical of the plexin family.

20 28. **Full-length PRO236 and PRO262 Polypeptides**

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO236 and PRO262. In particular, Applicants have identified and isolated cDNA encoding PRO236 and PRO262 polypeptides, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that various portions of the PRO236 and PRO262 polypeptides have significant homology with various β -galactosidase and β -galactosidase precursor polypeptides. Accordingly, it is presently believed that the PRO236 and PRO262 25 polypeptides disclosed in the present application are newly identified β -galactosidase homologs.

29. **Full-length PRO239 Polypeptides**

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO239. In particular, Applicants have identified and isolated cDNA 30 encoding a PRO239 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that various portions of the PRO239 polypeptide have significant homology with densin proteins. Accordingly, it is presently believed that PRO239 polypeptide disclosed in the present application is a newly identified member of the densin family and possesses cell adhesion 35 and the ability to effect synaptic processes as is typical of the densin family.

30. Full-length PRO257 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO257. In particular, Applicants have identified and isolated cDNA encoding a PRO257 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that various portions of the PRO257 polypeptide have significant homology with the ebnerin precursor and ebnerin protein. Accordingly, it is presently believed that PRO257 polypeptide disclosed in the present application is a newly identified protein member which is related to the ebnerin protein.

31. Full-length PRO260 Polypeptides

10 The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO260. In particular, Applicants have identified and isolated cDNA encoding a PRO260 polypeptide, as disclosed in further detail in the Examples below. Using programs such as BLAST and FastA sequence alignment computer programs, Applicants found that various portions of the PRO260 polypeptide have significant homology with the alpha-1-fucosidase precursor. Accordingly, it is presently believed that PRO260 polypeptide disclosed in the present application is a newly identified member 15 of the fucosidase family and possesses enzymatic activity related to fucose residues typical of the fucosidase family.

32. Full-length PRO263 Polypeptides

20 The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO263. In particular, Applicants have identified and isolated cDNA encoding a PRO263 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that various portions of the PRO263 polypeptide have significant homology with the CD44 antigen and related proteins. Accordingly, it is presently believed that 25 PRO263 polypeptide disclosed in the present application is a newly identified member of the CD44 antigen family and possesses at least one of the properties associated with these antigens, i.e., cancer and HIV marker, cell-cell or cell-matrix interactions, regulating cell traffic, lymph node homing, transmission of growth signals, and presentation of chemokines and growth factors to traveling cells.

30 33. Full-length PRO270 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO270. In particular, Applicants have identified and isolated cDNA encoding a PRO270 polypeptide, as disclosed in further detail in the Examples below. Using BLAST, FastA and sequence alignment computer programs, Applicants found that various portions of the PRO270 polypeptide have significant homology with various thioredoxin proteins. Accordingly, it is presently believed that PRO270 35 polypeptide disclosed in the present application is a newly identified member of the thioredoxin family and possesses the ability to effect reduction-oxidation (redox) state typical of the thioredoxin family.

34. Full-length PRO271 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO271. In particular, Applicants have identified and isolated cDNA encoding a PRO271 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that the PRO271 polypeptide has significant homology with various link proteins and precursors thereof. Accordingly, it is presently believed that PRO271 polypeptide disclosed in the present application is a newly identified link protein homolog.

35. Full-length PRO272 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO272. In particular, Applicants have identified and isolated cDNA encoding a PRO272 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that various portions of the PRO272 polypeptide have significant homology with the human reticulocalbin protein and its precursors. Applicants have also found that the DNA encoding the PRO272 polypeptide has significant homology with the mouse reticulocalbin precursor protein. Accordingly, it is presently believed that PRO272 polypeptide disclosed in the present application is a newly identified member of the reticulocalbin family and possesses the ability to bind calcium typical of the reticulocalbin family.

36. Full-length PRO294 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO294. In particular, Applicants have identified and isolated cDNA encoding a PRO294 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that various portions of the PRO294 polypeptide have significant homology with the various portions of a number of collagen proteins. Accordingly, it is presently believed that PRO294 polypeptide disclosed in the present application is a newly identified member of the collagen family.

37. Full-length PRO295 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO295. In particular, Applicants have identified and isolated cDNA encoding a PRO295 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that various portions of the PRO295 polypeptide have significant homology with integrin proteins. Accordingly, it is presently believed that PRO295 polypeptide disclosed in the present application is a newly identified member of the integrin family and possesses cell adhesion typical of the integrin family.

38. Full-length PRO293 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO293. In particular, Applicants have identified and isolated cDNA encoding a PRO293 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that portions of the PRO293 polypeptide have 5 significant homology with the neuronal leucine rich repeat proteins 1 and 2, (NLRR-1 and NLRR-2), particularly NLRR-2. Accordingly, it is presently believed that PRO293 polypeptide disclosed in the present application is a newly identified member of the neuronal leucine rich repeat protein family and possesses ligand-ligand binding activity typical of the NRLL protein family.

10 39. Full-length PRO247 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO247. In particular, Applicants have identified and isolated cDNA encoding a PRO247 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that various portions of the PRO247 polypeptide have 15 significant homology with densin. Applicants have also found that the DNA encoding the PRO247 polypeptide has significant homology with a number of other proteins, including KIAA0231. Accordingly, it is presently believed that PRO247 polypeptide disclosed in the present application is a newly identified member of the leucine rich repeat family and possesses ligand binding abilities typical of this family.

20 40. Full-length PRO302, PRO303, PRO304, PRO307 and PRO343 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO302, PRO303, PRO304, PRO307 and PRO343. In particular, Applicants have identified and isolated cDNA encoding PRO302, PRO303, PRO304, PRO307 and PRO343 polypeptides, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that various portions of the PRO302, PRO303, PRO304, PRO307 and 25 PRO343 polypeptides have significant homology with various protease proteins. Accordingly, it is presently believed that the PRO302, PRO303, PRO304, PRO307 and PRO343 polypeptides disclosed in the present application are newly identified protease proteins.

30 41. Full-length PRO328 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO328. In particular, Applicants have identified and isolated cDNA encoding a PRO328 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that various portions of the PRO328 polypeptide have 35 significant homology with the human glioblastoma protein ("GLIP"). Further, Applicants found that various portions of the PRO328 polypeptide have significant homology with the cysteine rich secretory protein ("CRISP") as identified by BLAST homology [ECCRISP3_1, S68683, and CRS3_HUMAN]. Accordingly, it

is presently believed that PRO328 polypeptide disclosed in the present application is a newly identified member of the GLIP or CRISP families and possesses transcriptional regulatory activity typical of the GLIP or CRISP families.

42. **Full-length PRO335, PRO331 and PRO326 Polypeptides**

5 The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO335, PRO331 or PRO326. In particular, Applicants have identified and isolated cDNA encoding a PRO335, PRO331 or PRO326 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that various portions of the PRO335, PRO331 or PRO326 polypeptide have significant homology with LIG-1, ALS 10 and in the case of PRO331, additionally, decorin. Accordingly, it is presently believed that the PRO335, PRO331 and PRO326 polypeptides disclosed in the present application are newly identified members of the leucine rich repeat superfamily, and particularly, are related to LIG-1 and possess the biological functions of this family as discussed and referenced herein.

15 43. **Full-length PRO332 Polypeptides**

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO332. In particular, Applicants have identified and isolated cDNA encoding PRO332 polypeptides, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that a full-length native sequence PRO332 (shown in 20 Figure 108 and SEQ ID NO:310) has about 30-40% amino acid sequence identity with a series of known proteoglycan sequences, including, for example, fibromodulin and fibromodulin precursor sequences of various species (FMOD_BOVIN, FMOD_CHICK, FMOD_RAT, FMOD_MOUSE, FMOD_HUMAN, P_R36773), osteomodulin sequences (AB000114_1, AB007848_1), decorin sequences (CFU83141_1, OCU03394_1, P_R42266, P_R42267, P_R42260, P_R89439), keratan sulfate proteoglycans (BTU48360_1, AF022890_1), corneal 25 proteoglycan (AF022256_1), and bone/cartilage proteoglycans and proteoglycane precursors (PGS1_BOVIN, PGS2_MOUSE, PGS2_HUMAN). Accordingly, it is presently believed that PRO332 disclosed in the present application is a new proteoglycan-type molecule, and may play a role in regulating extracellular matrix, cartilage, and/or bone function.

30 44. **Full-length PRO334 Polypeptides**

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO334. In particular, Applicants have identified and isolated cDNA encoding a PRO334 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that various portions of the PRO334 polypeptide have 35 significant homology with fibulin and fibrillin. Accordingly, it is presently believed that PRO334 polypeptide disclosed in the present application is a newly identified member of the epidermal growth factor family and possesses properties and activities typical of this family.

45. Full-length PRO346 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO346. In particular, Applicants have identified and isolated cDNA encoding a PRO346 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that a full-length native sequence PRO346 (shown in Figure 112 and SEQ ID NO:320) has 28 % amino acid sequence identity with carcinoembryonic antigen. 5 Accordingly, it is presently believed that PRO346 disclosed in the present application is a newly identified member of the carcinoembryonic protein family and may be expressed in association with neoplastic tissue disorders.

10 46. Full-length PRO268 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO268. In particular, Applicants have identified and isolated cDNA encoding a PRO268 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that portions of the PRO268 polypeptide have 15 significant homology with the various protein disulfide isomerase proteins. Accordingly, it is presently believed that PRO268 polypeptide disclosed in the present application is a homolog of the protein disulfide isomerase p5 protein.

15 47. Full-length PRO330 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO330. In particular, Applicants have identified and isolated cDNA encoding a PRO330 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that various portions of the PRO330 polypeptide have significant homology with the murine prolyl 4-hydroxylase alpha-II subunit protein. Accordingly, it is presently believed that PRO330 polypeptide disclosed in the present application is a novel prolyl 4-hydroxylase subunit 25 polypeptide.

20 48. Full-length PRO339 and PRO310 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO339 and PRO310. In particular, Applicants have identified and 30 isolated cDNA encoding a PRO339 polypeptide, as disclosed in further detail in the Examples below. Applicants have also identified and isolated cDNA encoding a PRO310 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that various portions of the PRO339 and PRO310 polypeptides have significant homology with small secreted 35 proteins from *C. elegans* and are distantly related to fringe. PRO339 also shows homology to collagen-like polymers. Sequences which were used to identify PRO310, designated herein as DNA40533 and DNA42267, also show homology to proteins from *C. elegans*. Accordingly, it is presently believed that the PRO339 and

PRO310 polypeptides disclosed in the present application are newly identified member of the family of proteins involved in development, and which may have regulatory abilities similar to the capability of fringe to regulate serrate.

49. Full Length PRO244 Polypeptides

5 The present invention provides newly identified and isolated nucleotide sequences encoding C-type lectins referred to in the present application as PRO244. In particular, applicants have identified and isolated cDNA encoding PRO244 polypeptides, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that a full-length native sequence PRO244 (shown in Figure 122 and SEQ ID NO:377) has 43% amino acid sequence identity with the hepatic lectin gallus gallus (LECH-CHICK), and 42% amino acid sequence identity with an HIV gp120 binding C-type lectin (A46274). Accordingly, it is presently believed that PRO244 disclosed in the present application is a newly identified member of the C-lectin superfamily and may play a role in immune function, apoptosis, or in the pathogenesis of atherosclerosis. In addition, PRO244 may be useful in identifying tumor-associated epitopes.

15 B. PRO Polypeptide Variants

In addition to the full-length native sequence PRO polypeptides described herein, it is contemplated that PRO variants can be prepared. PRO variants can be prepared by introducing appropriate nucleotide changes into the PRO DNA, and/or by synthesis of the desired PRO polypeptide. Those skilled in the art will appreciate that amino acid changes may alter post-translational processes of the PRO, such as changing the number or position 20 of glycosylation sites or altering the membrane anchoring characteristics.

Variations in the native full-length sequence PRO or in various domains of the PRO described herein, can be made, for example, using any of the techniques and guidelines for conservative and non-conservative mutations set forth, for instance, in U.S. Patent No. 5,364,934. Variations may be a substitution, deletion or insertion of one or more codons encoding the PRO that results in a change in the amino acid sequence of the 25 PRO as compared with the native sequence PRO. Optionally the variation is by substitution of at least one amino acid with any other amino acid in one or more of the domains of the PRO. Guidance in determining which amino acid residue may be inserted, substituted or deleted without adversely affecting the desired activity may be found by comparing the sequence of the PRO with that of homologous known protein molecules and minimizing the number of amino acid sequence changes made in regions of high homology. Amino acid 30 substitutions can be the result of replacing one amino acid with another amino acid having similar structural and/or chemical properties, such as the replacement of a leucine with a serine, i.e., conservative amino acid replacements. Insertions or deletions may optionally be in the range of about 1 to 5 amino acids. The variation allowed may be determined by systematically making insertions, deletions or substitutions of amino acids in the sequence and testing the resulting variants for activity exhibited by the full-length or mature native sequence.

35 PRO polypeptide fragments are provided herein. Such fragments may be truncated at the N-terminus or C-terminus, or may lack internal residues, for example, when compared with a full length native protein. Certain fragments lack amino acid residues that are not essential for a desired biological activity of the PRO

polypeptide.

PRO fragments may be prepared by any of a number of conventional techniques. Desired peptide fragments may be chemically synthesized. An alternative approach involves generating PRO fragments by enzymatic digestion, e.g., by treating the protein with an enzyme known to cleave proteins at sites defined by particular amino acid residues, or by digesting the DNA with suitable restriction enzymes and isolating the 5 desired fragment. Yet another suitable technique involves isolating and amplifying a DNA fragment encoding a desired polypeptide fragment, by polymerase chain reaction (PCR). Oligonucleotides that define the desired termini of the DNA fragment are employed at the 5' and 3' primers in the PCR. Preferably, PRO polypeptide fragments share at least one biological and/or immunological activity with the native PRO polypeptide disclosed herein.

10 In particular embodiments, conservative substitutions of interest are shown in Table 1 under the heading of preferred substitutions. If such substitutions result in a change in biological activity, then more substantial changes, denominated exemplary substitutions in Table 1, or as further described below in reference to amino acid classes, are introduced and the products screened.

Table 1

	<u>Original Residue</u>	<u>Exemplary Substitutions</u>	<u>Preferred Substitutions</u>
5	Ala (A)	val; leu; ile	val
	Arg (R)	lys; gln; asn	lys
	Asn (N)	gln; his; lys; arg	gln
	Asp (D)	glu	glu
	Cys (C)	ser	ser
10	Gln (Q)	asn	asn
	Glu (E)	asp	asp
	Gly (G)	pro; ala	ala
	His (H)	asn; gln; lys; arg	arg
	Ile (I)	leu; val; met; ala; phe; norleucine	leu
15	Leu (L)	norleucine; ile; val; met; ala; phe	ile
	Lys (K)	arg; gln; asn	arg
	Met (M)	leu; phe; ile	leu
20	Phe (F)	leu; val; ile; ala; tyr	leu
	Pro (P)	ala	ala
	Ser (S)	thr	thr
	Thr (T)	ser	ser
	Trp (W)	tyr; phe	tyr
25	Tyr (Y)	trp; phe; thr; ser	phe
	Val (V)	ile; leu; met; phe; ala; norleucine	leu

Substantial modifications in function or immunological identity of the PRO polypeptide are accomplished 30 by selecting substitutions that differ significantly in their effect on maintaining (a) the structure of the polypeptide backbone in the area of the substitution, for example, as a sheet or helical conformation, (b) the charge or hydrophobicity of the molecule at the target site, or (c) the bulk of the side chain. Naturally occurring residues are divided into groups based on common side-chain properties:

(1) hydrophobic: norleucine, met, ala, val, leu, ile;

35 (2) neutral hydrophilic: cys, ser, thr;

(3) acidic: asp, glu;

(4) basic: asn, gln, his, lys, arg;

(5) residues that influence chain orientation: gly, pro; and

(6) aromatic: trp, tyr, phe.

40 Non-conservative substitutions will entail exchanging a member of one of these classes for another class. Such substituted residues also may be introduced into the conservative substitution sites or, more preferably, into the remaining (non-conserved) sites.

The variations can be made using methods known in the art such as oligonucleotide-mediated (site-directed) mutagenesis, alanine scanning, and PCR mutagenesis. Site-directed mutagenesis [Carter et al., Nucl. Acids Res., 13:4331 (1986); Zoller et al., Nucl. Acids Res., 10:6487 (1987)], cassette mutagenesis [Wells et al., Gene, 34:315 (1985)], restriction selection mutagenesis [Wells et al., Philos. Trans. R. Soc. London Ser A, 317:415 (1986)] or other known techniques can be performed on the cloned DNA to produce the PRO variant

DNA.

Scanning amino acid analysis can also be employed to identify one or more amino acids along a contiguous sequence. Among the preferred scanning amino acids are relatively small, neutral amino acids.¹ Such amino acids include alanine, glycine, serine, and cysteine. Alanine is typically a preferred scanning amino acid among this group because it eliminates the side-chain beyond the beta-carbon and is less likely to alter the main-chain conformation of the variant [Cunningham and Wells, *Science*, 244: 1081-1085 (1989)]. Alanine is also typically preferred because it is the most common amino acid. Further, it is frequently found in both buried and exposed positions [Creighton, *The Proteins*, (W.H. Freeman & Co., N.Y.); Chothia, *J. Mol. Biol.*, 150:1 (1976)]. If alanine substitution does not yield adequate amounts of variant, an isoteric amino acid can be used.

10 C. Modifications of PRO

Covalent modifications of PRO are included within the scope of this invention. One type of covalent modification includes reacting targeted amino acid residues of a PRO polypeptide with an organic derivatizing agent that is capable of reacting with selected side chains or the N- or C-terminal residues of the PRO. Derivatization with bifunctional agents is useful, for instance, for crosslinking PRO to a water-insoluble support matrix or surface for use in the method for purifying anti-PRO antibodies, and vice-versa. Commonly used crosslinking agents include, e.g., 1,1-bis(diazoacetyl)-2-phenylethane, glutaraldehyde, N-hydroxysuccinimide esters, for example, esters with 4-azidosalicylic acid, homobifunctional imidoesters, including disuccinimidyl esters such as 3,3'-dithiobis(succinimidylpropionate), bifunctional maleimides such as bis-N-maleimido-1,8-octane and agents such as methyl-3-[(p-azidophenyl)dithio]propioimide.

20 Other modifications include deamidation of glutaminyl and asparaginyl residues to the corresponding glutamyl and aspartyl residues, respectively, hydroxylation of proline and lysine, phosphorylation of hydroxyl groups of seryl or threonyl residues, methylation of the α -amino groups of lysine, arginine, and histidine side chains [T.E. Creighton, *Proteins: Structure and Molecular Properties*, W.H. Freeman & Co., San Francisco, pp. 79-86 (1983)], acetylation of the N-terminal amine, and amidation of any C-terminal carboxyl group.

25 Another type of covalent modification of the PRO polypeptide included within the scope of this invention comprises altering the native glycosylation pattern of the polypeptide. "Altering the native glycosylation pattern" is intended for purposes herein to mean deleting one or more carbohydrate moieties found in native sequence PRO (either by removing the underlying glycosylation site or by deleting the glycosylation by chemical and/or enzymatic means), and/or adding one or more glycosylation sites that are not present in the 30 native sequence PRO. In addition, the phrase includes qualitative changes in the glycosylation of the native proteins, involving a change in the nature and proportions of the various carbohydrate moieties present.

35 Addition of glycosylation sites to the PRO polypeptide may be accomplished by altering the amino acid sequence. The alteration may be made, for example, by the addition of, or substitution by, one or more serine or threonine residues to the native sequence PRO (for O-linked glycosylation sites). The PRO amino acid sequence may optionally be altered through changes at the DNA level, particularly by mutating the DNA encoding the PRO polypeptide at preselected bases such that codons are generated that will translate into the desired amino acids.

Another means of increasing the number of carbohydrate moieties on the PRO polypeptide is by chemical or enzymatic coupling of glycosides to the polypeptide. Such methods are described in the art, e.g., in WO 87/05330 published 11 September 1987, and in Aplin and Wriston, CRC Crit. Rev. Biochem., pp. 259-306 (1981).

5 Removal of carbohydrate moieties present on the PRO polypeptide may be accomplished chemically or enzymatically or by mutational substitution of codons encoding for amino acid residues that serve as targets for glycosylation. Chemical deglycosylation techniques are known in the art and described, for instance, by Hakimuddin, et al., Arch. Biochem. Biophys., 259:52 (1987) and by Edge et al., Anal. Biochem., 118:131 (1981). Enzymatic cleavage of carbohydrate moieties on polypeptides can be achieved by the use of a variety of endo- and exo-glycosidases as described by Thotakura et al., Meth. Enzymol., 138:350 (1987).

10 Another type of covalent modification of PRO comprises linking the PRO polypeptide to one of a variety of nonproteinaceous polymers, e.g., polyethylene glycol (PEG), polypropylene glycol, or polyoxyalkylenes, in the manner set forth in U.S. Patent Nos. 4,640,835; 4,496,689; 4,301,144; 4,670,417; 4,791,192 or 4,179,337.

The PRO of the present invention may also be modified in a way to form a chimeric molecule comprising PRO fused to another, heterologous polypeptide or amino acid sequence.

15 In one embodiment, such a chimeric molecule comprises a fusion of the PRO with a tag polypeptide which provides an epitope to which an anti-tag antibody can selectively bind. The epitope tag is generally placed at the amino- or carboxyl- terminus of the PRO. The presence of such epitope-tagged forms of the PRO can be detected using an antibody against the tag polypeptide. Also, provision of the epitope tag enables the PRO to be readily purified by affinity purification using an anti-tag antibody or another type of affinity matrix that binds 20 to the epitope tag. Various tag polypeptides and their respective antibodies are well known in the art. Examples include poly-histidine (poly-his) or poly-histidine-glycine (poly-his-gly) tags; the flu HA tag polypeptide and its antibody 12CA5 [Field et al., Mol. Cell. Biol., 8:2159-2165 (1988)]; the c-myc tag and the 8F9, 3C7, 6E10, G4, B7 and 9E10 antibodies thereto [Evan et al., Molecular and Cellular Biology, 5:3610-3616 (1985)]; and the 25 Herpes Simplex virus glycoprotein D (gD) tag and its antibody [Paborsky et al., Protein Engineering, 3(6):547-553 (1990)]. Other tag polypeptides include the Flag-peptide [Hopp et al., BioTechnology, 6:1204-1210 (1988)]; the KT3 epitope peptide [Martin et al., Science, 255:192-194 (1992)]; an α -tubulin epitope peptide [Skinner et al., J. Biol. Chem., 266:15163-15166 (1991)]; and the T7 gene 10 protein peptide tag [Lutz-Freyermuth et al., Proc. Natl. Acad. Sci. USA, 87:6393-6397 (1990)].

30 In an alternative embodiment, the chimeric molecule may comprise a fusion of the PRO with an immunoglobulin or a particular region of an immunoglobulin. For a bivalent form of the chimeric molecule (also referred to as an "immunoadhesin"), such a fusion could be to the Fc region of an IgG molecule. The Ig fusions preferably include the substitution of a soluble (transmembrane domain deleted or inactivated) form of a PRO polypeptide in place of at least one variable region within an Ig molecule. In a particularly preferred embodiment, the immunoglobulin fusion includes the hinge, CH2 and CH3, or the hinge, CH1, CH2 and CH3 35 regions of an IgG1 molecule. For the production of immunoglobulin fusions see also US Patent No. 5,428,130 issued June 27, 1995.

5 D. Preparation of PRO

The description below relates primarily to production of PRO by culturing cells transformed or transfected with a vector containing PRO nucleic acid. It is, of course, contemplated that alternative methods, which are well known in the art, may be employed to prepare PRO. For instance, the PRO sequence, or portions thereof, may be produced by direct peptide synthesis using solid-phase techniques [see, e.g., Stewart et al., Solid-Phase Peptide Synthesis, W.H. Freeman Co., San Francisco, CA (1969); Merrifield, J. Am. Chem. Soc., 85:2149-2154 (1963)]. *In vitro* protein synthesis may be performed using manual techniques or by automation. Automated synthesis may be accomplished, for instance, using an Applied Biosystems Peptide Synthesizer (Foster City, CA) using manufacturer's instructions. Various portions of the PRO may be chemically synthesized separately and combined using chemical or enzymatic methods to produce the full-length

10 PRO.

15 1. Isolation of DNA Encoding PRO

DNA encoding PRO may be obtained from a cDNA library prepared from tissue believed to possess the PRO mRNA and to express it at a detectable level. Accordingly, human PRO DNA can be conveniently obtained from a cDNA library prepared from human tissue, such as described in the Examples. The PRO-encoding gene may also be obtained from a genomic library or by known synthetic procedures (e.g., automated nucleic acid synthesis).

20 Libraries can be screened with probes (such as antibodies to the PRO or oligonucleotides of at least about 20-80 bases) designed to identify the gene of interest or the protein encoded by it. Screening the cDNA or genomic library with the selected probe may be conducted using standard procedures, such as described in Sambrook et al., Molecular Cloning: A Laboratory Manual (New York: Cold Spring Harbor Laboratory Press, 1989). An alternative means to isolate the gene encoding PRO is to use PCR methodology [Sambrook et al., supra; Dieffenbach et al., PCR Primer: A Laboratory Manual (Cold Spring Harbor Laboratory Press, 1995)].

25 The Examples below describe techniques for screening a cDNA library. The oligonucleotide sequences selected as probes should be of sufficient length and sufficiently unambiguous that false positives are minimized. The oligonucleotide is preferably labeled such that it can be detected upon hybridization to DNA in the library being screened. Methods of labeling are well known in the art, and include the use of radiolabels like ³²P-labeled ATP, biotinylation or enzyme labeling. Hybridization conditions, including moderate stringency and high stringency, are provided in Sambrook et al., supra.

30 Sequences identified in such library screening methods can be compared and aligned to other known sequences deposited and available in public databases such as GenBank or other private sequence databases. Sequence identity (at either the amino acid or nucleotide level) within defined regions of the molecule or across the full-length sequence can be determined using methods known in the art and as described herein.

35 Nucleic acid having protein coding sequence may be obtained by screening selected cDNA or genomic libraries using the deduced amino acid sequence disclosed herein for the first time, and, if necessary, using conventional primer extension procedures as described in Sambrook et al., supra, to detect precursors and processing intermediates of mRNA that may not have been reverse-transcribed into cDNA.

2. Selection and Transformation of Host Cells

Host cells are transfected or transformed with expression or cloning vectors described herein for PRO production and cultured in conventional nutrient media modified as appropriate for inducing promoters, selecting transformants, or amplifying the genes encoding the desired sequences. The culture conditions, such as media, temperature, pH and the like, can be selected by the skilled artisan without undue experimentation. In general, 5 principles, protocols, and practical techniques for maximizing the productivity of cell cultures can be found in Mammalian Cell Biotechnology: a Practical Approach, M. Butler, ed. (IRL Press, 1991) and Sambrook et al., supra.

Methods of eukaryotic cell transfection and prokaryotic cell transformation are known to the ordinarily skilled artisan, for example, CaCl_2 , CaPO_4 , liposome-mediated and electroporation. Depending on the host cell 10 used, transformation is performed using standard techniques appropriate to such cells. The calcium treatment employing calcium chloride, as described in Sambrook et al., supra, or electroporation is generally used for prokaryotes. Infection with *Agrobacterium tumefaciens* is used for transformation of certain plant cells, as described by Shaw et al., Gene, 23:315 (1983) and WO 89/05859 published 29 June 1989. For mammalian cells without such cell walls, the calcium phosphate precipitation method of Graham and van der Eb, Virology, 15: 456-457 (1978) can be employed. General aspects of mammalian cell host system transfections have been 15 described in U.S. Patent No. 4,399,216. Transformations into yeast are typically carried out according to the method of Van Solingen et al., J. Bact., 130:946 (1977) and Hsiao et al., Proc. Natl. Acad. Sci. (USA), 76:3829 (1979). However, other methods for introducing DNA into cells, such as by nuclear microinjection, 20 electroporation, bacterial protoplast fusion with intact cells, or polycations, e.g., polybrene, polyornithine, may also be used. For various techniques for transforming mammalian cells, see Keown et al., Methods in Enzymology, 185:527-537 (1990) and Mansour et al., Nature, 336:348-352 (1988).

Suitable host cells for cloning or expressing the DNA in the vectors herein include prokaryote, yeast, or higher eukaryote cells. Suitable prokaryotes include but are not limited to eubacteria, such as Gram-negative or Gram-positive organisms, for example, Enterobacteriaceae such as *E. coli*. Various *E. coli* strains are 25 publicly available, such as *E. coli* K12 strain MM294 (ATCC 31,446); *E. coli* X1776 (ATCC 31,537); *E. coli* strain W3110 (ATCC 27,325) and K5 772 (ATCC 53,635). Other suitable prokaryotic host cells include Enterobacteriaceae such as *Escherichia*, e.g., *E. coli*, *Enterobacter*, *Erwinia*, *Klebsiella*, *Proteus*, *Salmonella*, e.g., *Salmonella typhimurium*, *Serratia*, e.g., *Serratia marcescens*, and *Shigella*, as well as *Bacilli* such as *B. subtilis* and *B. licheniformis* (e.g., *B. licheniformis* 41P disclosed in DD 266,710 published 12 April 1989). 30 *Pseudomonas* such as *P. aeruginosa*, and *Streptomyces*. These examples are illustrative rather than limiting. Strain W3110 is one particularly preferred host or parent host because it is a common host strain for recombinant DNA product fermentations. Preferably, the host cell secretes minimal amounts of proteolytic enzymes. For example, strain W3110 may be modified to effect a genetic mutation in the genes encoding proteins endogenous to the host, with examples of such hosts including *E. coli* W3110 strain 1A2, which has the complete genotype 35 *tonA*; *E. coli* W3110 strain 9E4, which has the complete genotype *tonA ptr3*; *E. coli* W3110 strain 27C7 (ATCC 55,244), which has the complete genotype *tonA ptr3 phoA E15 (argF-lac)169 degP ompT kan'*; *E. coli* W3110 strain 37D6, which has the complete genotype *tonA ptr3 phoA E15 (argF-lac)169 degP ompT rbs7*

ilvG kan^r; *E. coli* W3110 strain 40B4, which is strain 37D6 with a non-kanamycin resistant *degP* deletion mutation; and an *E. coli* strain having mutant periplasmic protease disclosed in U.S. Patent No. 4,946,783 issued 7 August 1990. Alternatively, *in vitro* methods of cloning, e.g., PCR or other nucleic acid polymerase reactions, are suitable.

In addition to prokaryotes, eukaryotic microbes such as filamentous fungi or yeast are suitable cloning 5 or expression hosts for PRO-encoding vectors. *Saccharomyces cerevisiae* is a commonly used lower eukaryotic host microorganism. Others include *Schizosaccharomyces pombe* (Beach and Nurse, *Nature*, 290: 140 [1981]; EP 139,383 published 2 May 1985); *Kluyveromyces* hosts (U.S. Patent No. 4,943,529; Fleer et al., *Bio/Technology*, 9:968-975 (1991)) such as, e.g., *K. lactis* (MW98-8C, CBS683, CBS4574; Louvencourt et al., *J. Bacteriol.*, 737 [1983]), *K. fragilis* (ATCC 12,424), *K. bulgaricus* (ATCC 16,045), *K. wickeramii* (ATCC 10 24,178), *K. waltii* (ATCC 56,500), *K. drosophilicola* (ATCC 36,906; Van den Berg et al., *Bio/Technology*, 8:135 (1990)), *K. thermotolerans*, and *K. marxianus*; *Yarrowia* (EP 402,226); *Pichia pastoris* (EP 183,070; Sreekrishna et al., *J. Basic Microbiol.*, 28:265-278 [1988]); *Candida*; *Trichoderma reesiae* (EP 244,234); *Neurospora crassa* (Case et al., *Proc. Natl. Acad. Sci. USA*, 76:5259-5263 [1979]); *Schwanniomyces* such as 15 *Schwanniomyces occidentalis* (EP 394,538 published 31 October 1990); and filamentous fungi such as, e.g., *Neurospora*, *Penicillium*, *Tolypocladium* (WO 91/00357 published 10 January 1991), and *Aspergillus* hosts such as *A. nidulans* (Ballance et al., *Biochem. Biophys. Res. Commun.*, 112:284-289 [1983]; Tilburn et al., *Gene*, 26:205-221 [1983]; Yelton et al., *Proc. Natl. Acad. Sci. USA*, 81: 1470-1474 [1984]) and *A. niger* (Kelly and Hynes, *EMBO J.*, 4:475-479 [1985]). Methylotrophic yeasts are suitable herein and include, but are not limited 20 to, yeast capable of growth on methanol selected from the genera consisting of *Hansenula*, *Candida*, *Kloeckera*, *Pichia*, *Saccharomyces*, *Torulopsis*, and *Rhodotorula*. A list of specific species that are exemplary of this class of yeasts may be found in C. Anthony, *The Biochemistry of Methylotrophs*, 269 (1982).

Suitable host cells for the expression of glycosylated PRO are derived from multicellular organisms. Examples of invertebrate cells include insect cells such as *Drosophila S2* and *Spodoptera Sf9*, as well as plant cells. Examples of useful mammalian host cell lines include Chinese hamster ovary (CHO) and COS cells. 25 More specific examples include monkey kidney CV1 line transformed by SV40 (COS-7, ATCC CRL 1651); human embryonic kidney line (293 or 293 cells subcloned for growth in suspension culture, Graham et al., *J. Gen Virol.*, 36:59 (1977)); Chinese hamster ovary cells/-DHFR (CHO, Urlaub and Chasin, *Proc. Natl. Acad. Sci. USA*, 77:4216 (1980)); mouse sertoli cells (TM4, Mather, *Biol. Reprod.*, 23:243-251 (1980)); human lung cells (W138, ATCC CCL 75); human liver cells (Hep G2, HB 8065); and mouse mammary tumor (MMT 30 060562, ATCC CCL51). The selection of the appropriate host cell is deemed to be within the skill in the art.

3. Selection and Use of a Replicable Vector

The nucleic acid (e.g., cDNA or genomic DNA) encoding PRO may be inserted into a replicable vector for cloning (amplification of the DNA) or for expression. Various vectors are publicly available. The vector 35 may, for example, be in the form of a plasmid, cosmid, viral particle, or phage. The appropriate nucleic acid sequence may be inserted into the vector by a variety of procedures. In general, DNA is inserted into an appropriate restriction endonuclease site(s) using techniques known in the art. Vector components generally

include, but are not limited to, one or more of a signal sequence, an origin of replication, one or more marker genes, an enhancer element, a promoter, and a transcription termination sequence. Construction of suitable vectors containing one or more of these components employs standard ligation techniques which are known to the skilled artisan.

The PRO may be produced recombinantly not only directly, but also as a fusion polypeptide with a 5 heterologous polypeptide, which may be a signal sequence or other polypeptide having a specific cleavage site at the N-terminus of the mature protein or polypeptide. In general, the signal sequence may be a component of the vector, or it may be a part of the PRO-encoding DNA that is inserted into the vector. The signal sequence may be a prokaryotic signal sequence selected, for example, from the group of the alkaline phosphatase, penicillinase, lpp, or heat-stable enterotoxin II leaders. For yeast secretion the signal sequence may be, e.g., 10 the yeast invertase leader, alpha factor leader (including *Saccharomyces* and *Kluyveromyces* α -factor leaders, the latter described in U.S. Patent No. 5,010,182), or acid phosphatase leader, the *C. albicans* glucoamylase leader (EP 362,179 published 4 April 1990), or the signal described in WO 90/13646 published 15 November 1990. In mammalian cell expression, mammalian signal sequences may be used to direct secretion of the 15 protein, such as signal sequences from secreted polypeptides of the same or related species, as well as viral secretory leaders.

Both expression and cloning vectors contain a nucleic acid sequence that enables the vector to replicate in one or more selected host cells. Such sequences are well known for a variety of bacteria, yeast, and viruses. The origin of replication from the plasmid pBR322 is suitable for most Gram-negative bacteria, the 2μ plasmid origin is suitable for yeast, and various viral origins (SV40, polyoma, adenovirus, VSV or BPV) are useful for 20 cloning vectors in mammalian cells.

Expression and cloning vectors will typically contain a selection gene, also termed a selectable marker. Typical selection genes encode proteins that (a) confer resistance to antibiotics or other toxins, e.g., ampicillin, neomycin, methotrexate, or tetracycline, (b) complement auxotrophic deficiencies, or (c) supply critical nutrients not available from complex media, e.g., the gene encoding D-alanine racemase for *Bacilli*.

25 An example of suitable selectable markers for mammalian cells are those that enable the identification of cells competent to take up the PRO-encoding nucleic acid, such as DHFR or thymidine kinase. An appropriate host cell when wild-type DHFR is employed is the CHO cell line deficient in DHFR activity, prepared and propagated as described by Urlaub et al., Proc. Natl. Acad. Sci. USA, 77:4216 (1980). A suitable selection gene for use in yeast is the *trp1* gene present in the yeast plasmid YRp7 [Stinchcomb et al., Nature, 282:39 (1979); Kingsman et al., Gene, 7:141 (1979); Tschemper et al., Gene, 10:157 (1980)]. The *trp1* gene 30 provides a selection marker for a mutant strain of yeast lacking the ability to grow in tryptophan, for example, ATCC No. 44076 or PEP4-1 [Jones, Genetics, 85:12 (1977)].

35 Expression and cloning vectors usually contain a promoter operably linked to the PRO-encoding nucleic acid sequence to direct mRNA synthesis. Promoters recognized by a variety of potential host cells are well known. Promoters suitable for use with prokaryotic hosts include the β -lactamase and lactose promoter systems [Chang et al., Nature, 275:615 (1978); Goeddel et al., Nature, 281:544 (1979)], alkaline phosphatase, a tryptophan (trp) promoter system [Goeddel, Nucleic Acids Res., 8:4057 (1980); EP 36,776], and hybrid

promoters such as the tac promoter [deBoer et al., Proc. Natl. Acad. Sci. USA, 80:21-25 (1983)]. Promoters for use in bacterial systems also will contain a Shine-Dalgarno (S.D.) sequence operably linked to the DNA encoding PRO.

Examples of suitable promoting sequences for use with yeast hosts include the promoters for 3-phosphoglycerate kinase [Hitzeman et al., J. Biol. Chem., 255:2073 (1980)] or other glycolytic enzymes [Hess 5 et al., J. Adv. Enzyme Reg., 7:149 (1968); Holland, Biochemistry, 17:4900 (1978)], such as enolase, glyceraldehyde-3-phosphate dehydrogenase, hexokinase, pyruvate decarboxylase, phosphofructokinase, glucose-6-phosphate isomerase, 3-phosphoglycerate mutase, pyruvate kinase, triosephosphateisomerase, phosphoglucose isomerase, and glucokinase.

Other yeast promoters, which are inducible promoters having the additional advantage of transcription 10 controlled by growth conditions, are the promoter regions for alcohol dehydrogenase 2, isocytochrome C, acid phosphatase, degradative enzymes associated with nitrogen metabolism, metallothionein, glyceraldehyde-3-phosphate dehydrogenase, and enzymes responsible for maltose and galactose utilization. Suitable vectors and promoters for use in yeast expression are further described in EP 73,657.

PRO transcription from vectors in mammalian host cells is controlled, for example, by promoters 15 obtained from the genomes of viruses such as polyoma virus, fowlpox virus (UK 2,211,504 published 5 July 1989), adenovirus (such as Adenovirus 2), bovine papilloma virus, avian sarcoma virus, cytomegalovirus, a retrovirus, hepatitis-B virus and Simian Virus 40 (SV40), from heterologous mammalian promoters, e.g., the actin promoter or an immunoglobulin promoter, and from heat-shock promoters, provided such promoters are compatible with the host cell systems.

20 Transcription of a DNA encoding the PRO by higher eukaryotes may be increased by inserting an enhancer sequence into the vector. Enhancers are cis-acting elements of DNA, usually about from 10 to 300 bp, that act on a promoter to increase its transcription. Many enhancer sequences are now known from mammalian genes (globin, elastase, albumin, α -fetoprotein, and insulin). Typically, however, one will use an enhancer from a eukaryotic cell virus. Examples include the SV40 enhancer on the late side of the replication 25 origin (bp 100-270), the cytomegalovirus early promoter enhancer, the polyoma enhancer on the late side of the replication origin, and adenovirus enhancers. The enhancer may be spliced into the vector at a position 5' or 3' to the PRO coding sequence, but is preferably located at a site 5' from the promoter.

Expression vectors used in eukaryotic host cells (yeast, fungi, insect, plant, animal, human, or nucleated 30 cells from other multicellular organisms) will also contain sequences necessary for the termination of transcription and for stabilizing the mRNA. Such sequences are commonly available from the 5' and, occasionally 3', untranslated regions of eukaryotic or viral DNAs or cDNAs. These regions contain nucleotide segments transcribed as polyadenylated fragments in the untranslated portion of the mRNA encoding PRO.

Still other methods, vectors, and host cells suitable for adaptation to the synthesis of PRO in recombinant vertebrate cell culture are described in Gething et al., Nature, 293:620-625 (1981); Mantei et al., 35 Nature, 281:40-46 (1979); EP 117,060; and EP 117,058.

4. Detecting Gene Amplification/Expression

Gene amplification and/or expression may be measured in a sample directly, for example, by conventional Southern blotting, Northern blotting to quantitate the transcription of mRNA [Thomas, Proc. Natl. Acad. Sci. USA, 77:5201-5205 (1980)], dot blotting (DNA analysis), or *in situ* hybridization, using an appropriately labeled probe, based on the sequences provided herein. Alternatively, antibodies may be employed 5 that can recognize specific duplexes, including DNA duplexes, RNA duplexes, and DNA-RNA hybrid duplexes or DNA-protein duplexes. The antibodies in turn may be labeled and the assay may be carried out where the duplex is bound to a surface, so that upon the formation of duplex on the surface, the presence of antibody bound to the duplex can be detected.

Gene expression, alternatively, may be measured by immunological methods, such as 10 immunohistochemical staining of cells or tissue sections and assay of cell culture or body fluids, to quantitate directly the expression of gene product. Antibodies useful for immunohistochemical staining and/or assay of sample fluids may be either monoclonal or polyclonal, and may be prepared in any mammal. Conveniently, the antibodies may be prepared against a native sequence PRO polypeptide or against a synthetic peptide based on the DNA sequences provided herein or against exogenous sequence fused to PRO DNA and encoding a specific 15 antibody epitope.

5. Purification of Polypeptide

Forms of PRO may be recovered from culture medium or from host cell lysates. If membrane-bound, it can be released from the membrane using a suitable detergent solution (e.g. Triton-X 100) or by enzymatic 20 cleavage. Cells employed in expression of PRO can be disrupted by various physical or chemical means, such as freeze-thaw cycling, sonication, mechanical disruption, or cell lysing agents.

It may be desired to purify PRO from recombinant cell proteins or polypeptides. The following procedures are exemplary of suitable purification procedures: by fractionation on an ion-exchange column; ethanol precipitation; reverse phase HPLC; chromatography on silica or on a cation-exchange resin such as 25 DEAE; chromatofocusing; SDS-PAGE; ammonium sulfate precipitation; gel filtration using, for example, Sephadex G-75; protein A Sepharose columns to remove contaminants such as IgG; and metal chelating columns to bind epitope-tagged forms of the PRO. Various methods of protein purification may be employed and such methods are known in the art and described for example in Deutscher, Methods in Enzymology, 182 (1990); Scopes, Protein Purification: Principles and Practice, Springer-Verlag, New York (1982). The purification 30 step(s) selected will depend, for example, on the nature of the production process used and the particular PRO produced.

E. Uses for PRO

Nucleotide sequences (or their complement) encoding PRO have various applications in the art of 35 molecular biology, including uses as hybridization probes, in chromosome and gene mapping and in the generation of anti-sense RNA and DNA. PRO nucleic acid will also be useful for the preparation of PRO polypeptides by the recombinant techniques described herein.

The full-length native sequence PRO gene, or portions thereof, may be used as hybridization probes for a cDNA library to isolate the full-length PRO cDNA or to isolate still other cDNAs (for instance, those encoding naturally-occurring variants of PRO or PRO from other species) which have a desired sequence identity to the native PRO sequence disclosed herein. Optionally, the length of the probes will be about 20 to about 50 bases. The hybridization probes may be derived from at least partially novel regions of the full length native 5 nucleotide sequence wherein those regions may be determined without undue experimentation or from genomic sequences including promoters, enhancer elements and introns of native sequence PRO. By way of example, a screening method will comprise isolating the coding region of the PRO gene using the known DNA sequence to synthesize a selected probe of about 40 bases. Hybridization probes may be labeled by a variety of labels, including radionucleotides such as ^{32}P or ^{35}S , or enzymatic labels such as alkaline phosphatase coupled to the 10 probe via avidin/biotin coupling systems. Labeled probes having a sequence complementary to that of the PRO gene of the present invention can be used to screen libraries of human cDNA, genomic DNA or mRNA to determine which members of such libraries the probe hybridizes to. Hybridization techniques are described in further detail in the Examples below.

Any EST sequences disclosed in the present application may similarly be employed as probes, using 15 the methods disclosed herein.

Other useful fragments of the PRO nucleic acids include antisense or sense oligonucleotides comprising a single-stranded nucleic acid sequence (either RNA or DNA) capable of binding to target PRO mRNA (sense) or PRO DNA (antisense) sequences. Antisense or sense oligonucleotides, according to the present invention, comprise a fragment of the coding region of PRO DNA. Such a fragment generally comprises at least about 14 20 nucleotides, preferably from about 14 to 30 nucleotides. The ability to derive an antisense or a sense oligonucleotide, based upon a cDNA sequence encoding a given protein is described in, for example, Stein and Cohen (*Cancer Res.* 48:2659, 1988) and van der Krol et al. (*BioTechniques* 6:958, 1988).

Binding of antisense or sense oligonucleotides to target nucleic acid sequences results in the formation of duplexes that block transcription or translation of the target sequence by one of several means, including 25 enhanced degradation of the duplexes, premature termination of transcription or translation, or by other means. The antisense oligonucleotides thus may be used to block expression of PRO proteins. Antisense or sense oligonucleotides further comprise oligonucleotides having modified sugar-phosphodiester backbones (or other sugar linkages, such as those described in WO 91/06629) and wherein such sugar linkages are resistant to endogenous nucleases. Such oligonucleotides with resistant sugar linkages are stable *in vivo* (i.e., capable of 30 resisting enzymatic degradation) but retain sequence specificity to be able to bind to target nucleotide sequences.

Other examples of sense or antisense oligonucleotides include those oligonucleotides which are covalently linked to organic moieties, such as those described in WO 90/10048, and other moieties that increases affinity of the oligonucleotide for a target nucleic acid sequence, such as poly-(L-lysine). Further still, intercalating agents, such as ellipticine, and alkylating agents or metal complexes may be attached to sense or 35 antisense oligonucleotides to modify binding specificities of the antisense or sense oligonucleotide for the target nucleotide sequence.

Antisense or sense oligonucleotides may be introduced into a cell containing the target nucleic acid

sequence by any gene transfer method, including, for example, CaPO_4 -mediated DNA transfection, electroporation, or by using gene transfer vectors such as Epstein-Barr virus. In a preferred procedure, an antisense or sense oligonucleotide is inserted into a suitable retroviral vector. A cell containing the target nucleic acid sequence is contacted with the recombinant retroviral vector, either *in vivo* or *ex vivo*. Suitable retroviral vectors include, but are not limited to, those derived from the murine retrovirus M-MuLV, N2 (a retrovirus 5 derived from M-MuLV), or the double copy vectors designated DCT5A, DCT5B and DCT5C (see WO 90/13641).

Sense or antisense oligonucleotides also may be introduced into a cell containing the target nucleotide sequence by formation of a conjugate with a ligand binding molecule, as described in WO 91/04753. Suitable ligand binding molecules include, but are not limited to, cell surface receptors, growth factors, other cytokines, 10 or other ligands that bind to cell surface receptors. Preferably, conjugation of the ligand binding molecule does not substantially interfere with the ability of the ligand binding molecule to bind to its corresponding molecule or receptor, or block entry of the sense or antisense oligonucleotide or its conjugated version into the cell.

Alternatively, a sense or an antisense oligonucleotide may be introduced into a cell containing the target nucleic acid sequence by formation of an oligonucleotide-lipid complex, as described in WO 90/10448. The 15 sense or antisense oligonucleotide-lipid complex is preferably dissociated within the cell by an endogenous lipase.

Antisense RNA or DNA molecules are generally at least about 5 bases in length, about 10 bases in length, about 15 bases in length, about 20 bases in length, about 25 bases in length, about 30 bases in length, about 35 bases in length, about 40 bases in length, about 45 bases in length, about 50 bases in length, about 55 bases in length, about 60 bases in length, about 65 bases in length, about 70 bases in length, about 75 bases in 20 length, about 80 bases in length, about 85 bases in length, about 90 bases in length, about 95 bases in length, about 100 bases in length, or more.

The probes may also be employed in PCR techniques to generate a pool of sequences for identification of closely related PRO coding sequences.

Nucleotide sequences encoding a PRO can also be used to construct hybridization probes for mapping 25 the gene which encodes that PRO and for the genetic analysis of individuals with genetic disorders. The nucleotide sequences provided herein may be mapped to a chromosome and specific regions of a chromosome using known techniques, such as *in situ* hybridization, linkage analysis against known chromosomal markers, and hybridization screening with libraries.

When the coding sequences for PRO encode a protein which binds to another protein (example, where 30 the PRO is a receptor), the PRO can be used in assays to identify the other proteins or molecules involved in the binding interaction. By such methods, inhibitors of the receptor/ligand binding interaction can be identified. Proteins involved in such binding interactions can also be used to screen for peptide or small molecule inhibitors or agonists of the binding interaction. Also, the receptor PRO can be used to isolate correlative ligand(s). Screening assays can be designed to find lead compounds that mimic the biological activity of a native PRO or 35 a receptor for PRO. Such screening assays will include assays amenable to high-throughput screening of chemical libraries, making them particularly suitable for identifying small molecule drug candidates. Small molecules contemplated include synthetic organic or inorganic compounds. The assays can be performed in a

variety of formats, including protein-protein binding assays, biochemical screening assays, immunoassays and cell based assays, which are well characterized in the art.

Nucleic acids which encode PRO or its modified forms can also be used to generate either transgenic animals or "knock out" animals which, in turn, are useful in the development and screening of therapeutically useful reagents. A transgenic animal (e.g., a mouse or rat) is an animal having cells that contain a transgene, 5 which transgene was introduced into the animal or an ancestor of the animal at a prenatal, e.g., an embryonic stage. A transgene is a DNA which is integrated into the genome of a cell from which a transgenic animal develops. In one embodiment, cDNA encoding PRO can be used to clone genomic DNA encoding PRO in accordance with established techniques and the genomic sequences used to generate transgenic animals that contain cells which express DNA encoding PRO. Methods for generating transgenic animals, particularly 10 animals such as mice or rats, have become conventional in the art and are described, for example, in U.S. Patent Nos. 4,736,866 and 4,870,009. Typically, particular cells would be targeted for PRO transgene incorporation with tissue-specific enhancers. Transgenic animals that include a copy of a transgene encoding PRO introduced into the germ line of the animal at an embryonic stage can be used to examine the effect of increased expression 15 of DNA encoding PRO. Such animals can be used as tester animals for reagents thought to confer protection from, for example, pathological conditions associated with its overexpression. In accordance with this facet of the invention, an animal is treated with the reagent and a reduced incidence of the pathological condition, compared to untreated animals bearing the transgene, would indicate a potential therapeutic intervention for the pathological condition.

Alternatively, non-human homologues of PRO can be used to construct a PRO "knock out" animal 20 which has a defective or altered gene encoding PRO as a result of homologous recombination between the endogenous gene encoding PRO and altered genomic DNA encoding PRO introduced into an embryonic stem cell of the animal. For example, cDNA encoding PRO can be used to clone genomic DNA encoding PRO in accordance with established techniques. A portion of the genomic DNA encoding PRO can be deleted or replaced with another gene, such as a gene encoding a selectable marker which can be used to monitor 25 integration. Typically, several kilobases of unaltered flanking DNA (both at the 5' and 3' ends) are included in the vector [see e.g., Thomas and Capecchi, *Cell*, 51:503 (1987) for a description of homologous recombination vectors]. The vector is introduced into an embryonic stem cell line (e.g., by electroporation) and cells in which the introduced DNA has homologously recombined with the endogenous DNA are selected [see e.g., Li et al., *Cell*, 69:915 (1992)]. The selected cells are then injected into a blastocyst of an animal (e.g., 30 a mouse or rat) to form aggregation chimeras [see e.g., Bradley, in *Teratocarcinomas and Embryonic Stem Cells: A Practical Approach*, E. J. Robertson, ed. (IRL, Oxford, 1987), pp. 113-152]. A chimeric embryo can then be implanted into a suitable pseudopregnant female foster animal and the embryo brought to term to create 35 a "knock out" animal. Progeny harboring the homologously recombined DNA in their germ cells can be identified by standard techniques and used to breed animals in which all cells of the animal contain the homologously recombined DNA. Knockout animals can be characterized for instance, for their ability to defend against certain pathological conditions and for their development of pathological conditions due to absence of the PRO polypeptide.

Nucleic acid encoding the PRO polypeptides may also be used in gene therapy. In gene therapy applications, genes are introduced into cells in order to achieve *in vivo* synthesis of a therapeutically effective genetic product, for example for replacement of a defective gene. "Gene therapy" includes both conventional gene therapy where a lasting effect is achieved by a single treatment, and the administration of gene therapeutic agents, which involves the one time or repeated administration of a therapeutically effective DNA or mRNA.

5 Antisense RNAs and DNAs can be used as therapeutic agents for blocking the expression of certain genes *in vivo*. It has already been shown that short antisense oligonucleotides can be imported into cells where they act as inhibitors, despite their low intracellular concentrations caused by their restricted uptake by the cell membrane. (Zamecnik *et al.*, Proc. Natl. Acad. Sci. USA 83:4143-4146 [1986]). The oligonucleotides can be modified to enhance their uptake, e.g. by substituting their negatively charged phosphodiester groups by
10 uncharged groups.

There are a variety of techniques available for introducing nucleic acids into viable cells. The techniques vary depending upon whether the nucleic acid is transferred into cultured cells *in vitro*, or *in vivo* in the cells of the intended host. Techniques suitable for the transfer of nucleic acid into mammalian cells *in vitro* include the use of liposomes, electroporation, microinjection, cell fusion, DEAE-dextran, the calcium phosphate
15 precipitation method, etc. The currently preferred *in vivo* gene transfer techniques include transfection with viral (typically retroviral) vectors and viral coat protein-liposome mediated transfection (Dzau *et al.*, Trends in Biotechnology 11, 205-210 [1993]). In some situations it is desirable to provide the nucleic acid source with an agent that targets the target cells, such as an antibody specific for a cell surface membrane protein or the target cell, a ligand for a receptor on the target cell, etc. Where liposomes are employed, proteins which bind
20 to a cell surface membrane protein associated with endocytosis may be used for targeting and/or to facilitate uptake, e.g. capsid proteins or fragments thereof tropic for a particular cell type, antibodies for proteins which undergo internalization in cycling, proteins that target intracellular localization and enhance intracellular half-life. The technique of receptor-mediated endocytosis is described, for example, by Wu *et al.*, J. Biol. Chem. 262, 4429-4432 (1987); and Wagner *et al.*, Proc. Natl. Acad. Sci. USA 87, 3410-3414 (1990). For review of gene
25 marking and gene therapy protocols see Anderson *et al.*, Science 256, 808-813 (1992).

The PRO polypeptides described herein may also be employed as molecular weight markers for protein electrophoresis purposes and the isolated nucleic acid sequences may be used for recombinantly expressing those markers.

The nucleic acid molecules encoding the PRO polypeptides or fragments thereof described herein are
30 useful for chromosome identification. In this regard, there exists an ongoing need to identify new chromosome markers, since relatively few chromosome marking reagents, based upon actual sequence data are presently available. Each PRO nucleic acid molecule of the present invention can be used as a chromosome marker.

The PRO polypeptides and nucleic acid molecules of the present invention may also be used for tissue typing, wherein the PRO polypeptides of the present invention may be differentially expressed in one tissue as
35 compared to another. PRO nucleic acid molecules will find use for generating probes for PCR, Northern analysis, Southern analysis and Western analysis.

The PRO polypeptides described herein may also be employed as therapeutic agents. The PRO polypeptides of the present invention can be formulated according to known methods to prepare pharmaceutically useful compositions, whereby the PRO product hereof is combined in admixture with a pharmaceutically acceptable carrier vehicle. Therapeutic formulations are prepared for storage by mixing the active ingredient having the desired degree of purity with optional physiologically acceptable carriers, excipients or stabilizers 5 (Remington's Pharmaceutical Sciences 16th edition, Osol, A. Ed. (1980)), in the form of lyophilized formulations or aqueous solutions. Acceptable carriers, excipients or stabilizers are nontoxic to recipients at the dosages and concentrations employed, and include buffers such as phosphate, citrate and other organic acids; antioxidants including ascorbic acid; low molecular weight (less than about 10 residues) polypeptides; proteins, such as serum albumin, gelatin or immunoglobulins; hydrophilic polymers such as polyvinylpyrrolidone, amino 10 acids such as glycine, glutamine, asparagine, arginine or lysine; monosaccharides, disaccharides and other carbohydrates including glucose, mannose, or dextrans; chelating agents such as EDTA; sugar alcohols such as mannitol or sorbitol; salt-forming counterions such as sodium; and/or nonionic surfactants such as TWEEN™, PLURONICS™ or PEG.

15 The formulations to be used for *in vivo* administration must be sterile. This is readily accomplished by filtration through sterile filtration membranes, prior to or following lyophilization and reconstitution.

Therapeutic compositions herein generally are placed into a container having a sterile access port, for example, an intravenous solution bag or vial having a stopper pierceable by a hypodermic injection needle.

20 The route of administration is in accord with known methods, e.g. injection or infusion by intravenous, intraperitoneal, intracerebral, intramuscular, intraocular, intraarterial or intralesional routes, topical administration, or by sustained release systems.

Dosages and desired drug concentrations of pharmaceutical compositions of the present invention may vary depending on the particular use envisioned. The determination of the appropriate dosage or route of 25 administration is well within the skill of an ordinary physician. Animal experiments provide reliable guidance for the determination of effective doses for human therapy. Interspecies scaling of effective doses can be performed following the principles laid down by Mordenti, J. and Chappell, W. "The use of interspecies scaling in toxicokinetics" In *Toxicokinetics and New Drug Development*, Yacobi et al., Eds., Pergamon Press, New York 1989, pp. 42-96.

When *in vivo* administration of a PRO polypeptide or agonist or antagonist thereof is employed, normal dosage amounts may vary from about 10 ng/kg to up to 100 mg/kg of mammal body weight or more per day, 30 preferably about 1 μ g/kg/day to 10 mg/kg/day, depending upon the route of administration. Guidance as to particular dosages and methods of delivery is provided in the literature; see, for example, U.S. Pat. Nos. 4,657,760; 5,206,344; or 5,225,212. It is anticipated that different formulations will be effective for different treatment compounds and different disorders, that administration targeting one organ or tissue, for example, may necessitate delivery in a manner different from that to another organ or tissue.

35 Where sustained-release administration of a PRO polypeptide is desired in a formulation with release characteristics suitable for the treatment of any disease or disorder requiring administration of the PRO polypeptide, microencapsulation of the PRO polypeptide is contemplated. Microencapsulation of recombinant

proteins for sustained release has been successfully performed with human growth hormone (rhGH), interferon- (rhIFN-), interleukin-2, and MN rgp120. Johnson et al., Nat. Med., 2:795-799 (1996); Yasuda, Biomed. Ther., 27:1221-1223 (1993); Hora et al., Bio/Technology, 8:755-758 (1990); Cleland, "Design and Production of Single Immunization Vaccines Using Polylactide Polyglycolide Microsphere Systems," in Vaccine Design: The Subunit and Adjuvant Approach, Powell and Newman, eds, (Plenum Press: New York, 1995), pp. 439-462;

5 WO 97/03692, WO 96/40072, WO 96/07399; and U.S. Pat. No. 5,654,010.

The sustained-release formulations of these proteins were developed using poly-lactic-coglycolic acid (PLGA) polymer due to its biocompatibility and wide range of biodegradable properties. The degradation products of PLGA, lactic and glycolic acids, can be cleared quickly within the human body. Moreover, the degradability of this polymer can be adjusted from months to years depending on its molecular weight and 10 composition. Lewis, "Controlled release of bioactive agents from lactide/glycolide polymer," in: M. Chasin and R. Langer (Eds.), Biodegradable Polymers as Drug Delivery Systems (Marcel Dekker: New York, 1990), pp. 1-41.

This invention encompasses methods of screening compounds to identify those that mimic the PRO polypeptide (agonists) or prevent the effect of the PRO polypeptide (antagonists). Screening assays for 15 antagonist drug candidates are designed to identify compounds that bind or complex with the PRO polypeptides encoded by the genes identified herein, or otherwise interfere with the interaction of the encoded polypeptides with other cellular proteins. Such screening assays will include assays amenable to high-throughput screening of chemical libraries, making them particularly suitable for identifying small molecule drug candidates.

The assays can be performed in a variety of formats, including protein-protein binding assays, 20 biochemical screening assays, immunoassays, and cell-based assays, which are well characterized in the art.

All assays for antagonists are common in that they call for contacting the drug candidate with a PRO polypeptide encoded by a nucleic acid identified herein under conditions and for a time sufficient to allow these two components to interact.

In binding assays, the interaction is binding and the complex formed can be isolated or detected in the 25 reaction mixture. In a particular embodiment, the PRO polypeptide encoded by the gene identified herein or the drug candidate is immobilized on a solid phase, e.g., on a microtiter plate, by covalent or non-covalent attachments. Non-covalent attachment generally is accomplished by coating the solid surface with a solution of the PRO polypeptide and drying. Alternatively, an immobilized antibody, e.g., a monoclonal antibody, specific for the PRO polypeptide to be immobilized can be used to anchor it to a solid surface. The assay is performed 30 by adding the non-immobilized component, which may be labeled by a detectable label, to the immobilized component, e.g., the coated surface containing the anchored component. When the reaction is complete, the non-reacted components are removed, e.g., by washing, and complexes anchored on the solid surface are detected. When the originally non-immobilized component carries a detectable label, the detection of label 35 immobilized on the surface indicates that complexing occurred. Where the originally non-immobilized component does not carry a label, complexing can be detected, for example, by using a labeled antibody specifically binding the immobilized complex.

If the candidate compound interacts with but does not bind to a particular PRO polypeptide encoded by a gene identified herein, its interaction with that polypeptide can be assayed by methods well known for detecting protein-protein interactions. Such assays include traditional approaches, such as, e.g., cross-linking, co-immunoprecipitation, and co-purification through gradients or chromatographic columns. In addition, protein-protein interactions can be monitored by using a yeast-based genetic system described by Fields and co-workers

5 (Fields and Song, Nature (London), 340:245-246 (1989); Chien et al., Proc. Natl. Acad. Sci. USA, 88:9578-9582 (1991)) as disclosed by Chevray and Nathans, Proc. Natl. Acad. Sci. USA, 89: 5789-5793 (1991). Many transcriptional activators, such as yeast GAL4, consist of two physically discrete modular domains, one acting as the DNA-binding domain, the other one functioning as the transcription-activation domain. The yeast expression system described in the foregoing publications (generally referred to as the "two-hybrid system")

10 takes advantage of this property, and employs two hybrid proteins, one in which the target protein is fused to the DNA-binding domain of GAL4, and another, in which candidate activating proteins are fused to the activation domain. The expression of a GAL1-lacZ reporter gene under control of a GAL4-activated promoter depends on reconstitution of GAL4 activity via protein-protein interaction. Colonies containing interacting polypeptides are detected with a chromogenic substrate for β -galactosidase. A complete kit

15 (MATCHMAKERTM) for identifying protein-protein interactions between two specific proteins using the two-hybrid technique is commercially available from Clontech. This system can also be extended to map protein domains involved in specific protein interactions as well as to pinpoint amino acid residues that are crucial for these interactions.

Compounds that interfere with the interaction of a gene encoding a PRO polypeptide identified herein

20 and other intra- or extracellular components can be tested as follows: usually a reaction mixture is prepared containing the product of the gene and the intra- or extracellular component under conditions and for a time allowing for the interaction and binding of the two products. To test the ability of a candidate compound to inhibit binding, the reaction is run in the absence and in the presence of the test compound. In addition, a placebo may be added to a third reaction mixture, to serve as positive control. The binding (complex formation)

25 between the test compound and the intra- or extracellular component present in the mixture is monitored as described hereinabove. The formation of a complex in the control reaction(s) but not in the reaction mixture containing the test compound indicates that the test compound interferes with the interaction of the test compound and its reaction partner.

To assay for antagonists, the PRO polypeptide may be added to a cell along with the compound to be

30 screened for a particular activity and the ability of the compound to inhibit the activity of interest in the presence of the PRO polypeptide indicates that the compound is an antagonist to the PRO polypeptide. Alternatively, antagonists may be detected by combining the PRO polypeptide and a potential antagonist with membrane-bound PRO polypeptide receptors or recombinant receptors under appropriate conditions for a competitive inhibition assay. The PRO polypeptide can be labeled, such as by radioactivity, such that the number of PRO polypeptide

35 molecules bound to the receptor can be used to determine the effectiveness of the potential antagonist. The gene encoding the receptor can be identified by numerous methods known to those of skill in the art, for example, ligand panning and FACS sorting. Coligan et al., Current Protocols in Immun., 1(2): Chapter 5 (1991).

Preferably, expression cloning is employed wherein polyadenylated RNA is prepared from a cell responsive to the PRO polypeptide and a cDNA library created from this RNA is divided into pools and used to transfect COS cells or other cells that are not responsive to the PRO polypeptide. Transfected cells that are grown on glass slides are exposed to labeled PRO polypeptide. The PRO polypeptide can be labeled by a variety of means including iodination or inclusion of a recognition site for a site-specific protein kinase. Following fixation and 5 incubation, the slides are subjected to autoradiographic analysis. Positive pools are identified and sub-pools are prepared and re-transfected using an interactive sub-pooling and re-screening process, eventually yielding a single clone that encodes the putative receptor.

As an alternative approach for receptor identification, labeled PRO polypeptide can be photoaffinity-linked with cell membrane or extract preparations that express the receptor molecule. Cross-linked material is 10 resolved by PAGE and exposed to X-ray film. The labeled complex containing the receptor can be excised, resolved into peptide fragments, and subjected to protein micro-sequencing. The amino acid sequence obtained from micro- sequencing would be used to design a set of degenerate oligonucleotide probes to screen a cDNA library to identify the gene encoding the putative receptor.

In another assay for antagonists, mammalian cells or a membrane preparation expressing the receptor 15 would be incubated with labeled PRO polypeptide in the presence of the candidate compound. The ability of the compound to enhance or block this interaction could then be measured.

More specific examples of potential antagonists include an oligonucleotide that binds to the fusions of immunoglobulin with PRO polypeptide, and, in particular, antibodies including, without limitation, poly- and 20 monoclonal antibodies and antibody fragments, single-chain antibodies, anti-idiotypic antibodies, and chimeric or humanized versions of such antibodies or fragments, as well as human antibodies and antibody fragments. Alternatively, a potential antagonist may be a closely related protein, for example, a mutated form of the PRO 25 polypeptide that recognizes the receptor but imparts no effect, thereby competitively inhibiting the action of the PRO polypeptide.

Another potential PRO polypeptide antagonist is an antisense RNA or DNA construct prepared using 30 antisense technology, where, e.g., an antisense RNA or DNA molecule acts to block directly the translation of mRNA by hybridizing to targeted mRNA and preventing protein translation. Antisense technology can be used to control gene expression through triple-helix formation or antisense DNA or RNA, both of which methods are based on binding of a polynucleotide to DNA or RNA. For example, the 5' coding portion of the polynucleotide sequence, which encodes the mature PRO polypeptides herein, is used to design an antisense RNA 35 oligonucleotide of from about 10 to 40 base pairs in length. A DNA oligonucleotide is designed to be complementary to a region of the gene involved in transcription (triple helix - see Lee et al., Nucl. Acids Res., 6:3073 (1979); Cooney et al., Science, 241: 456 (1988); Dervan et al., Science, 251:1360 (1991)), thereby preventing transcription and the production of the PRO polypeptide. The antisense RNA oligonucleotide hybridizes to the mRNA *in vivo* and blocks translation of the mRNA molecule into the PRO polypeptide 40 (antisense - Okano, Neurochem., 56:560 (1991); Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression (CRC Press: Boca Raton, FL, 1988). The oligonucleotides described above can also be delivered to cells such that the antisense RNA or DNA may be expressed *in vivo* to inhibit production of the PRO 45

polypeptide. When antisense DNA is used, oligodeoxyribonucleotides derived from the translation-initiation site, e.g., between about -10 and +10 positions of the target gene nucleotide sequence, are preferred.

Potential antagonists include small molecules that bind to the active site, the receptor binding site, or growth factor or other relevant binding site of the PRO polypeptide, thereby blocking the normal biological activity of the PRO polypeptide. Examples of small molecules include, but are not limited to, small peptides or peptide-like molecules, preferably soluble peptides, and synthetic non-peptidyl organic or inorganic compounds.

Ribozymes are enzymatic RNA molecules capable of catalyzing the specific cleavage of RNA. Ribozymes act by sequence-specific hybridization to the complementary target RNA, followed by endonucleolytic cleavage. Specific ribozyme cleavage sites within a potential RNA target can be identified by known techniques.

10 For further details see, e.g., Rossi, *Current Biology*, 4:469-471 (1994), and PCT publication No. WO 97/33551 (published September 18, 1997).

Nucleic acid molecules in triple-helix formation used to inhibit transcription should be single-stranded and composed of deoxynucleotides. The base composition of these oligonucleotides is designed such that it promotes triple-helix formation via Hoogsteen base-pairing rules, which generally require sizeable stretches of 15 purines or pyrimidines on one strand of a duplex. For further details see, e.g., PCT publication No. WO 97/33551, *supra*.

These small molecules can be identified by any one or more of the screening assays discussed hereinabove and/or by any other screening techniques well known for those skilled in the art.

With regard to the PRO211 and PRO217 polypeptide, therapeutic indications include disorders 20 associated with the preservation and maintenance of gastrointestinal mucosa and the repair of acute and chronic mucosal lesions (e.g., enterocolitis, Zollinger-Ellison syndrome, gastrointestinal ulceration and congenital microvillus atrophy), skin diseases associated with abnormal keratinocyte differentiation (e.g., psoriasis, epithelial cancers such as lung squamous cell carcinoma, epidermoid carcinoma of the vulva and gliomas).

Since the PRO232 polypeptide and nucleic acid encoding it possess sequence homology to a cell surface 25 stem cell antigen and its encoding nucleic acid, probes based upon the PRO232 nucleotide sequence may be employed to identify other novel stem cell surface antigen proteins. Soluble forms of the PRO232 polypeptide may be employed as antagonists of membrane bound PRO232 activity both *in vitro* and *in vivo*. PRO232 polypeptides may be employed in screening assays designed to identify agonists or antagonists of the native 30 PRO232 polypeptide, wherein such assays may take the form of any conventional cell-type or biochemical binding assay. Moreover, the PRO232 polypeptide may serve as a molecular marker for the tissues in which the polypeptide is specifically expressed.

With regard to the PRO187 polypeptides disclosed herein, FGF-8 has been implicated in cellular 35 differentiation and embryogenesis, including the patterning which appears during limb formation. FGF-8 and the PRO187 molecules of the invention therefore are likely to have potent effects on cell growth and development. Diseases which relate to cellular growth and differentiation are therefore suitable targets for therapeutics based on functionality similar to FGF-8. For example, diseases related to growth or survival of nerve cells including Parkinson's disease, Alzheimer's disease, ALS, neuropathies. Additionally, disease related

to uncontrolled cell growth, e.g., cancer, would also be expected therapeutic targets.

With regard to the PRO265 polypeptides disclosed herein, other methods for use with PRO265 are described in U.S. Patent 5,654,270 to Ruoslahti et al. In particular, PRO265 can be used in comparison with the fibromodulin disclosed therein to compare its effects on reducing dermal scarring and other properties of the fibromodulin described therein including where it is located and with what it binds and does not.

5 The PRO219 polypeptides of the present invention which play a regulatory role in the blood coagulation cascade may be employed *in vivo* for therapeutic purposes as well as for *in vitro* purposes. Those of ordinary skill in the art will well know how to employ PRO219 polypeptides for such uses.

10 The PRO246 polypeptides of the present invention which serve as cell surface receptors for one or more viruses will find other uses. For example, extracellular domains derived from these PRO246 polypeptides may be employed therapeutically *in vivo* for lessening the effects of viral infection. Those PRO246 polypeptides which serve as tumor specific antigens may be exploited as therapeutic targets for anti-tumor drugs, and the like. Those of ordinary skill in the art will well know how to employ PRO246 polypeptides for such uses.

15 Assays in which connective growth factor and other growth factors are usually used should be performed with PRO261. An assay to determine whether TGF beta induces PRO261, indicating a role in cancer is performed as known in the art. Wound repair and tissue growth assays are also performed with PRO261. The results are applied accordingly.

20 PRO228 polypeptides should be used in assays in which EMR1, CD97 and latrophilin would be used in to determine their relative activities. The results can be applied accordingly. For example, a competitive binding assay with PRO228 and CD97 can be performed with the ligand for CD97, CD55.

25 Native PRO533 is a 216 amino acid polypeptide of which residues 1-22 are the signal sequence. Residues 3 to 216 have a Blast score of 509, corresponding to 53% homology to fibroblast growth factor. At the nucleotide level, DNA47412, the EST from which PCR oligos were generated to isolate the full length DNA49435-1219, has been observed to map to 11p15. Sequence homology to the 11p15 locus would indicate that PRO533 may have utility in the treatment of Usher Syndrome or Atrophy areata.

30 As mentioned previously, fibroblast growth factors can act upon cells in both a mitogenic and non-mitogenic manner. These factors are mitogenic for a wide variety of normal diploid mesoderm-derived and neural crest-derived cells, inducing granulosa cells, adrenal cortical cells, chondrocytes, myoblasts, corneal and vascular endothelial cells (bovine or human), vascular smooth muscle cells, lens, retina and prostatic epithelial cells, oligodendrocytes, astrocytes, chondrocytes, myoblasts and osteoblasts.

35 Non-mitogenic actions of fibroblast growth factors include promotion of cell migration into a wound area (chemotaxis), initiation of new blood vessel formulation (angiogenesis), modulation of nerve regeneration and survival (neurotrophism), modulation of endocrine functions, and stimulation or suppression of specific cellular protein expression, extracellular matrix production and cell survival. Baird, A. & Bohlen, P., *Handbook of Exp. Pharmacol.* 95(1): 369-418 (1990). These properties provide a basis for using fibroblast growth factors in therapeutic approaches to accelerate wound healing, nerve repair, collateral blood vessel formation, and the like. For example, fibroblast growth factors, have been suggested to minimize myocardium damage in heart disease and surgery (U.S.P. 4,378,437).

Since the PRO245 polypeptide and nucleic acid encoding it possess sequence homology to a transmembrane protein tyrosine kinase protein and its encoding nucleic acid, probes based upon the PRO245 nucleotide sequence may be employed to identify other novel transmembrane tyrosine kinase proteins. Soluble forms of the PRO245 polypeptide may be employed as antagonists of membrane bound PRO245 activity both *in vitro* and *in vivo*. PRO245 polypeptides may be employed in screening assays designed to identify agonists 5 or antagonists of the native PRO245 polypeptide, wherein such assays may take the form of any conventional cell-type or biochemical binding assay. Moreover, the PRO245 polypeptide may serve as a molecular marker for the tissues in which the polypeptide is specifically expressed.

PRO220, PRO221 and PRO227 all have leucine rich repeats. Additionally, PRO220 and PRO221 have homology to SLIT and leucine rich repeat protein. Therefore, these proteins are useful in assays described in 10 the literature, *supra*, wherein the SLIT and leucine rich repeat protein are used. Regarding the SLIT protein, PRO227 can be used in an assay to determine the affect of PRO227 on neurodegenerative disease. Additionally, PRO227 has homology to human glycoprotein V. In the case of PRO227, this polypeptide is used in an assay to determine its affect on bleeding, clotting, tissue repair and scarring.

The PRO266 polypeptide can be used in assays to determine if it has a role in neurodegenerative 15 diseases or their reversal.

PRO269 polypeptides and portions thereof which effect the activity of thrombin may also be useful for *in vivo* therapeutic purposes, as well as for various *in vitro* applications. In addition, PRO269 polypeptides and portions thereof may have therapeutic use as an antithrombotic agent with reduced risk for hemorrhage as compared with heparin. Peptides having homology to thrombomodulin are particularly desirable.

20 PRO287 polypeptides and portions thereof which effect the activity of bone morphogenic protein "BMP1"/procollagen C-proteinase (PCP) may also be useful for *in vivo* therapeutic purposes, as well as for various *in vitro* applications. In addition, PRO287 polypeptides and portions thereof may have therapeutic applications in wound healing and tissue repair. Peptides having homology to procollagen C-proteinase enhancer protein and its precursor may also be used to induce bone and/or cartilage formation and are therefore of 25 particular interest to the scientific and medical communities.

Therapeutic indications for PRO214 polypeptides include disorders associated with the preservation and maintenance of gastrointestinal mucosa and the repair of acute and chronic mucosal lesions (e.g., enterocolitis, Zollinger-Ellison syndrome, gastrointestinal ulceration and congenital microvillus atrophy), skin diseases associated with abnormal keratinocyte differentiation (e.g., psoriasis, epithelial cancers such as lung squamous 30 cell carcinoma, epidermoid carcinoma of the vulva and gliomas).

Studies on the generation and analysis of mice deficient in members of the TGF- superfamily are reported in Matzuk, *Trends in Endocrinol. and Metabol.*, 6: 120-127 (1995).

35 The PRO317 polypeptide, as well as PRO317-specific antibodies, inhibitors, agonists, receptors, or their analogs, herein are useful in treating PRO317-associated disorders. Hence, for example, they may be employed in modulating endometrial bleeding angiogenesis, and may also have an effect on kidney tissue. Endometrial bleeding can occur in gynecological diseases such as endometrial cancer as abnormal bleeding. Thus, the compositions herein may find use in diagnosing and treating abnormal bleeding conditions in the endometrium,

as by reducing or eliminating the need for a hysterectomy. The molecules herein may also find use in angiogenesis applications such as anti-tumor indications for which the antibody against vascular endothelial growth factor is used, or, conversely, ischemic indications for which vascular endothelial growth factor is employed.

5 Bioactive compositions comprising PRO317 or agonists or antagonists thereof may be administered in a suitable therapeutic dose determined by any of several methodologies including clinical studies on mammalian species to determine maximal tolerable dose and on normal human subjects to determine safe dose. Additionally, the bioactive agent may be complexed with a variety of well established compounds or compositions which enhance stability or pharmacological properties such as half-life. It is contemplated that the therapeutic, bioactive composition may be delivered by intravenous infusion into the bloodstream or any other effective means which 10 could be used for treating problems of the kidney, uterus, endometrium, blood vessels, or related tissue, e.g., in the heart or genital tract.

Dosages and administration of PRO317, PRO317 agonist, or PRO317 antagonist in a pharmaceutical composition may be determined by one of ordinary skill in the art of clinical pharmacology or pharmacokinetics. See, for example, Mordenti and Rescigno, Pharmaceutical Research, 9:17-25 (1992); Morenti *et al.*, 15 Pharmaceutical Research, 8:1351-1359 (1991); and Mordenti and Chappell, "The use of interspecies scaling in toxicokinetics" in Toxicokinetics and New Drug Development, Yacobi *et al.* (eds) (Pergamon Press: NY, 1989), pp. 42-96. An effective amount of PRO317, PRO317 agonist, or PRO317 antagonist to be employed therapeutically will depend, for example, upon the therapeutic objectives, the route of administration, and the condition of the mammal. Accordingly, it will be necessary for the therapist to titer the dosage and modify the 20 route of administration as required to obtain the optimal therapeutic effect. A typical daily dosage might range from about 10 ng/kg to up to 100 mg/kg of the mammal's body weight or more per day, preferably about 1 μ g/kg/day to 10 mg/kg/day. Typically, the clinician will administer PRO317, PRO317 agonist, or PRO317 antagonist, until a dosage is reached that achieves the desired effect for treatment of the above mentioned disorders.

25 PRO317 or an PRO317 agonist or PRO317 antagonist may be administered alone or in combination with another to achieve the desired pharmacological effect. PRO317 itself, or agonists or antagonists of PRO317 can provide different effects when administered therapeutically. Such compounds for treatment will be formulated in a nontoxic, inert, pharmaceutically acceptable aqueous carrier medium preferably at a pH of about 5 to 8, more preferably 6 to 8, although the pH may vary according to the characteristics of the PRO317, agonist, or 30 antagonist being formulated and the condition to be treated. Characteristics of the treatment compounds include solubility of the molecule, half-life, and antigenicity/immunogenicity; these and other characteristics may aid in defining an effective carrier.

PRO317 or PRO317 agonists or PRO317 antagonists may be delivered by known routes of administration including but not limited to topical creams and gels; transmucosal spray and aerosol, transdermal 35 patch and bandage; injectable, intravenous, and lavage formulations; and orally administered liquids and pills, particularly formulated to resist stomach acid and enzymes. The particular formulation, exact dosage, and route of administration will be determined by the attending physician and will vary according to each specific situation.

Such determinations of administration are made by considering multiple variables such as the condition to be treated, the type of mammal to be treated, the compound to be administered, and the pharmacokinetic profile of the particular treatment compound. Additional factors which may be taken into account include disease state (e.g. severity) of the patient, age, weight, gender, diet, time of administration, drug combination, reaction sensitivities, and tolerance/response to therapy. Long-acting treatment compound formulations (such as 5 liposomally encapsulated PRO317 or PEGylated PRO317 or PRO317 polymeric microspheres, such as polylactic acid-based microspheres) might be administered every 3 to 4 days, every week, or once every two weeks depending on half-life and clearance rate of the particular treatment compound.

Normal dosage amounts may vary from about 10 ng/kg to up to 100 mg/kg of mammal body weight or more per day, preferably about 1 μ g/kg/day to 10 mg/kg/day, depending upon the route of administration. 10 Guidance as to particular dosages and methods of delivery is provided in the literature; see, for example, U.S. Pat. Nos. 4,657,760; 5,206,344; or 5,225,212. It is anticipated that different formulations will be effective for different treatment compounds and different disorders, that administration targeting the uterus, for example, may necessitate delivery in a manner different from that to another organ or tissue, such as cardiac tissue.

Where sustained-release administration of PRO317 is desired in a formulation with release 15 characteristics suitable for the treatment of any disease or disorder requiring administration of PRO317, microencapsulation of PRO317 is contemplated. Microencapsulation of recombinant proteins for sustained release has been successfully performed with human growth hormone (rhGH), interferon- (rhIFN-), interleukin-2, and MN rgp120. Johnson *et al.*, Nat. Med., 2: 795-799 (1996); Yasuda, Biomed. Ther., 27: 1221-1223 (1993); Hora *et al.*, Bio/Technology, 8: 755-758 (1990); Cleland, "Design and Production of Single 20 Immunization Vaccines Using Polylactide Polyglycolide Microsphere Systems," in Vaccine Design: The Subunit and Adjuvant Approach, Powell and Newman, eds, (Plenum Press: New York, 1995), pp. 439-462; WO 97/03692, WO 96/40072, WO 96/07399; and U.S Pat. No. 5,654,010.

It is contemplated that conditions or diseases of the uterus, endometrial tissue, or other genital tissues or cardiac tissues may precipitate damage that is treatable with PRO317 or PRO317 agonist where PRO317 25 expression is reduced in the diseased state; or with antibodies to PRO317 or other PRO317 antagonists where the expression of PRO317 is increased in the diseased state. These conditions or diseases may be specifically diagnosed by the probing tests discussed above for physiologic and pathologic problems which affect the function of the organ.

The PRO317, PRO317 agonist, or PRO317 antagonist may be administered to a mammal with another 30 biologically active agent, either separately or in the same formulation to treat a common indication for which they are appropriate. For example, it is contemplated that PRO317 can be administered together with EBAF-1 for those indications on which they demonstrate the same qualitative biological effects. Alternatively, where they have opposite effects, EBAF-1 may be administered together with an antagonist to PRO317, such as an anti-PRO317 antibody. Further, PRO317 may be administered together with VEGF for coronary ischemia where 35 such indication is warranted, or with an anti-VEGF for cancer as warranted, or, conversely, an antagonist to PRO317 may be administered with VEGF for coronary ischemia or with anti-VEGF to treat cancer as warranted. These administrations would be in effective amounts for treating such disorders.

Native PRO301 (SEQ ID NO:119) has a Blast score of 246 and 30% homology at residues 24 to 282 of Figure 44 with A33_HUMAN, an A33 antigen precursor. A33 antigen precursor, as explained in the Background is a tumor-specific antigen, and as such, is a recognized marker and therapeutic target for the diagnosis and treatment of colon cancer. The expression of tumor-specific antigens is often associated with the progression of neoplastic tissue disorders. Native PRO301 (SEQ ID NO:119) and A33_HUMAN also show a 5 Blast score of 245 and 30% homology at residues 21 to 282 of Fig. 44 with A33_HUMAN, the variation dependent upon how spaces are inserted into the compared sequences. Native PRO301 (SEQ ID NO:119) also has a Blast score of 165 and 29% homology at residues 60 to 255 of Fig. 44 with HS46KDA_1, a human coxsackie and adenovirus receptor protein, also known as cell surface protein HCAR. This region of PRO301 also shows a similar Blast score and homology with HSU90716_1. Expression of such proteins is usually 10 associated with viral infection and therapeutics for the prevention of such infection may be accordingly conceived. As mentioned in the Background, the expression of viral receptors is often associated with neoplastic tumors.

Therapeutic uses for the PRO234 polypeptides of the invention includes treatments associated with leukocyte homing or the interaction between leukocytes and the endothelium during an inflammatory response. 15 Examples include asthma, rheumatoid arthritis, psoriasis and multiple sclerosis.

Since the PRO231 polypeptide and nucleic acid encoding it possess sequence homology to a putative acid phosphatase and its encoding nucleic acid, probes based upon the PRO231 nucleotide sequence may be employed to identify other novel phosphatase proteins. Soluble forms of the PRO231 polypeptide may be employed as antagonists of membrane bound PRO231 activity both *in vitro* and *in vivo*. PRO231 polypeptides 20 may be employed in screening assays designed to identify agonists or antagonists of the native PRO231 polypeptide, wherein such assays may take the form of any conventional cell-type or biochemical binding assay. Moreover, the PRO231 polypeptide may serve as a molecular marker for the tissues in which the polypeptide is specifically expressed.

PRO229 polypeptides can be fused with peptides of interest to determine whether the fusion peptide has 25 an increased half-life over the peptide of interest. The PRO229 polypeptides can be used accordingly to increase the half-life of polypeptides of interest. Portions of PRO229 which cause the increase in half-life are an embodiment of the invention herein.

PRO238 can be used in assays which measure its ability to reduce substrates, including oxygen and Aceyl-CoA, and particularly, measure PRO238's ability to produce oxygen free radicals. This is done by using 30 assays which have been previously described. PRO238 can further be used to assay for candidates which block, reduce or reverse its reducing abilities. This is done by performing side by side assays where candidates are added in one assay having PRO238 and a substrate to reduce, and not added in another assay, being the same but for the lack of the presence of the candidate.

PRO233 polypeptides and portions thereof which have homology to reductase may also be useful for 35 *in vivo* therapeutic purposes, as well as for various other applications. The identification of novel reductase proteins and related molecules may be relevant to a number of human disorders such as inflammatory disease, organ failure, atherosclerosis, cardiac injury, infertility, birth defects, premature aging, AIDS, cancer, diabetic

complications and mutations in general. Given that oxygen free radicals and antioxidants appear to play important roles in a number of disease processes, the identification of new reductase proteins and reductase-like molecules is of special importance in that such proteins may serve as potential therapeutics for a variety of different human disorders. Such polypeptides may also play important roles in biotechnological and medical research, as well as various industrial applications. As a result, there is particular scientific and medical interest 5 in new molecules, such as PRO233.

The PRO223 polypeptides of the present invention which exhibit serine carboxypeptidase activity may be employed *in vivo* for therapeutic purposes as well as for *in vitro* purposes. Those of ordinary skill in the art will well know how to employ PRO223 polypeptides for such uses.

10 PRO235 polypeptides and portions thereof which may be involved in cell adhesion are also useful for *in vivo* therapeutic purposes, as well as for various *in vitro* applications. In addition, PRO235 polypeptides and portions thereof may have therapeutic applications in disease states which involve cell adhesion. Given the physiological importance of cell adhesion mechanisms *in vivo*, efforts are currently being undertaken to identify new, native proteins which are involved in cell adhesion. Therefore, peptides having homology to plexin are of particular interest to the scientific and medical communities.

15 Because the PRO236 and PRO262 polypeptides disclosed herein are homologous to various known β -galactosidase proteins, the PRO236 and PRO262 polypeptides disclosed herein will find use in conjugates of monoclonal antibodies and the polypeptide for specific killing of tumor cells by generation of active drug from a galactosylated prodrug (e.g., the generation of 5-fluorouridine from the prodrug β -D-galactosyl-5-fluorouridine). The PRO236 and PRO262 polypeptides disclosed herein may also find various uses both *in vivo* 20 and *in vitro*, wherein those uses will be similar or identical to uses for which β -galactosidase proteins are now employed. Those of ordinary skill in the art will well know how to employ PRO236 and PRO262 polypeptides for such uses.

25 PRO239 polypeptides and portions thereof which have homology to densin may also be useful for *in vivo* therapeutic purposes, as well as for various *in vitro* applications. In addition, PRO239 polypeptides and portions thereof may have therapeutic applications in disease states which involve synaptic mechanisms, regeneration or cell adhesion. Given the physiological importance of synaptic processes, regeneration and cell adhesion mechanisms *in vivo*, efforts are currently being undertaken to identify new, native proteins which are involved in synaptic machinery and cell adhesion. Therefore, peptides having homology to densin are of particular interest to the scientific and medical communities.

30 The PRO260 polypeptides described herein can be used in assays to determine their relation to fucosidase. In particular, the PRO260 polypeptides can be used in assays in determining their ability to remove fucose or other sugar residues from proteoglycans. The PRO260 polypeptides can be assayed to determine if they have any functional or locational similarities as fucosidase. The PRO260 polypeptides can then be used to regulate the systems in which they are integral.

35 PRO263 can be used in assays wherein CD44 antigen is generally used to determine PRO263 activity relative to that of CD44. The results can be used accordingly.

PRO270 polypeptides and portions thereof which effect reduction-oxidation (redox) state may also be

useful for *in vivo* therapeutic purposes, as well as for various *in vitro* applications. More specifically, PRO270 polypeptides may affect the expression of a large variety of genes thought to be involved in the pathogenesis of AIDS, cancer, atherosclerosis, diabetic complications and in pathological conditions involving oxidative stress such as stroke and inflammation. In addition, PRO270 polypeptides and portions thereof may affect the expression of genes which have a role in apoptosis. Therefore, peptides having homology to thioredoxin are 5 particularly desirable to the scientific and medical communities.

PRO272 polypeptides and portions thereof which possess the ability to bind calcium may also have numerous *in vivo* therapeutic uses, as well as various *in vitro* applications. Therefore, peptides having homology to reticulocalbin are particularly desirable. Those with ordinary skill in the art will know how to employ PRO272 polypeptides and portions thereof for such purposes.

10 PRO294 polypeptides and portions thereof which have homology to collagen may also be useful for *in vivo* therapeutic purposes, as well as for various other applications. The identification of novel collagens and collagen-like molecules may have relevance to a number of human disorders. Thus, the identification of new collagens and collagen-like molecules is of special importance in that such proteins may serve as potential therapeutics for a variety of different human disorders. Such polypeptides may also play important roles in 15 biotechnological and medical research as well as various industrial applications. Given the large number of uses for collagen, there is substantial interest in polypeptides with homology to the collagen molecule.

20 PRO295 polypeptides and portions thereof which have homology to integrin may also be useful for *in vivo* therapeutic purposes, as well as for various other applications. The identification of novel integrins and integrin-like molecules may have relevance to a number of human disorders such as modulating the binding or activity of cells of the immune system. Thus, the identification of new integrins and integrin-like molecules is of special importance in that such proteins may serve as potential therapeutics for a variety of different human disorders. Such polypeptides may also play important roles in biotechnological and medical research as well as various industrial applications. As a result, there is particular scientific and medical interest in new molecules, such as PRO295.

25 As the PRO293 polypeptide is clearly a leucine rich repeat polypeptide homologue, the peptide can be used in all applications that the known NLRR-1 and NLRR-2 polypeptides are used. The activity can be compared between these peptides and thus applied accordingly.

30 The PRO247 polypeptides described herein can be used in assays in which densin is used to determine the activity of PRO247 relative to densin or these other proteins. The results can be used accordingly in diagnostics and/or therapeutic applications with PRO247.

PRO302, PRO303, PRO304, PRO307 and PRO343 polypeptides of the present invention which possess protease activity may be employed both *in vivo* for therapeutic purposes and *in vitro*. Those of ordinary skill in the art will well know how to employ the PRO302, PRO303, PRO304, PRO307 and PRO343 polypeptides of the present invention for such purposes.

35 PRO328 polypeptides and portions thereof which have homology to GLIP and CRISP may also be useful for *in vivo* therapeutic purposes, as well as for various other applications. The identification of novel GLIP and CRISP-like molecules may have relevance to a number of human disorders which involve transcriptional

regulation or are over expressed in human tumors. Thus, the identification of new GLIP and CRISP-like molecules is of special importance in that such proteins may serve as potential therapeutics for a variety of different human disorders. Such polypeptides may also play important roles in biotechnological and medical research as well as in various industrial applications. As a result, there is particular scientific and medical interest in new molecules, such as PRO328.

5 Uses for PRO335, PRO331 or PRO326 including uses in competitive assays with LIG-1, ALS and decorin to determine their relative activities. The results can be used accordingly. PRO335, PRO331 or PRO326 can also be used in assays where LIG-1 would be used to determine if the same effects are incurred.

PRO332 contains GAG repeat (GKEK) at amino acid positions 625-628 in Fig. 108 (SEQ ID NO:310). Slippage in such repeats can be associated with human disease. Accordingly, PRO332 can be useful for the 10 treatment of such disease conditions by gene therapy, i.e. by introduction of a gene containing the correct GKEK sequence motif.

Other uses of PRO334 include use in assays in which fibrillin or fibulin would be used to determine the relative activity of PRO334 to fibrillin or fibulin. In particular, PRO334 can be used in assays which require the mechanisms imparted by epidermal growth factor repeats.

15 Native PRO346 (SEQ ID NO:320) has a Blast score of 230, corresponding to 27% homology between amino acid residues 21 to 343 with residues 35 to 1040 CGM6_HUMAN, a carcinoembryonic antigen cgm6 precursor. This homology region includes nearly all but 2 N-terminal extracellular domain residues, including an immunoglobulin superfamily homology at residues 148 to 339 of PRO346 in addition to several transmembrane residues (340-343). Carcinoembryonic antigen precursor, as explained in the Background is a 20 tumor-specific antigen, and as such, is a recognized marker and therapeutic target for the diagnosis and treatment of colon cancer. The expression of tumor-specific antigens is often associated with the progression of neoplastic tissue disorders. Native PRO346 (SEQ ID NO:320) and P_W06874, a human carcinoembryonic antigen CEA-d have a Blast score of 224 and homology of 28% between residues 2 to 343 and 67 to 342, respectively. This homology includes the entire extracellular domain residues of native PRO346, minus the initiator methionine 25 (residues 2 to 18) as well as several transmembrane residues (340-343).

PRO268 polypeptides which have protein disulfide isomerase activity will be useful for many applications where protein disulfide isomerase activity is desirable including, for example, for use in promoting proper disulfide bond formation in recombinantly produced proteins so as to increase the yield of correctly folded protein. Those of ordinary skill in the art will readily know how to employ such PRO268 polypeptides for such 30 purposes.

PRO330 polypeptides of the present invention which possess biological activity related to that of the prolyl 4-hydroxylase alpha subunit protein may be employed both *in vivo* for therapeutic purposes and *in vitro*. Those of ordinary skill in the art will well know how to employ the PRO330 polypeptides of the present invention for such purposes.

F. Anti-PRO Antibodies

The present invention further provides anti-PRO antibodies. Exemplary antibodies include polyclonal, monoclonal, humanized, bispecific, and heteroconjugate antibodies.

1. Polyclonal Antibodies

5 The anti-PRO antibodies may comprise polyclonal antibodies. Methods of preparing polyclonal antibodies are known to the skilled artisan. Polyclonal antibodies can be raised in a mammal, for example, by one or more injections of an immunizing agent and, if desired, an adjuvant. Typically, the immunizing agent and/or adjuvant will be injected in the mammal by multiple subcutaneous or intraperitoneal injections. The immunizing agent may include the PRO polypeptide or a fusion protein thereof. It may be useful to conjugate 10 the immunizing agent to a protein known to be immunogenic in the mammal being immunized. Examples of such immunogenic proteins include but are not limited to keyhole limpet hemocyanin, serum albumin, bovine thyroglobulin, and soybean trypsin inhibitor. Examples of adjuvants which may be employed include Freund's complete adjuvant and MPL-TDM adjuvant (monophosphoryl Lipid A, synthetic trehalose dicorynomycolate). The immunization protocol may be selected by one skilled in the art without undue experimentation.

15

2. Monoclonal Antibodies

The anti-PRO antibodies may, alternatively, be monoclonal antibodies. Monoclonal antibodies may be prepared using hybridoma methods, such as those described by Kohler and Milstein, *Nature*, 256:495 (1975). In a hybridoma method, a mouse, hamster, or other appropriate host animal, is typically immunized with an 20 immunizing agent to elicit lymphocytes that produce or are capable of producing antibodies that will specifically bind to the immunizing agent. Alternatively, the lymphocytes may be immunized *in vitro*.

The immunizing agent will typically include the PRO polypeptide or a fusion protein thereof. Generally, either peripheral blood lymphocytes ("PBLs") are used if cells of human origin are desired, or spleen cells or lymph node cells are used if non-human mammalian sources are desired. The lymphocytes are then 25 fused with an immortalized cell line using a suitable fusing agent, such as polyethylene glycol, to form a hybridoma cell [Goding, *Monoclonal Antibodies: Principles and Practice*, Academic Press, (1986) pp. 59-103]. Immortalized cell lines are usually transformed mammalian cells, particularly myeloma cells of rodent, bovine and human origin. Usually, rat or mouse myeloma cell lines are employed. The hybridoma cells may be cultured in a suitable culture medium that preferably contains one or more substances that inhibit the growth or 30 survival of the unfused, immortalized cells. For example, if the parental cells lack the enzyme hypoxanthine guanine phosphoribosyl transferase (HGPRT or HPRT), the culture medium for the hybridomas typically will include hypoxanthine, aminopterin, and thymidine ("HAT medium"), which substances prevent the growth of HGPRT-deficient cells.

Preferred immortalized cell lines are those that fuse efficiently, support stable high level expression of 35 antibody by the selected antibody-producing cells, and are sensitive to a medium such as HAT medium. More preferred immortalized cell lines are murine myeloma lines, which can be obtained, for instance, from the Salk Institute Cell Distribution Center, San Diego, California and the American Type Culture Collection, Manassas,

Virginia. Human myeloma and mouse-human heteromyeloma cell lines also have been described for the production of human monoclonal antibodies [Kozbor, J. Immunol., 133:3001 (1984); Brodeur et al., Monoclonal Antibody Production Techniques and Applications, Marcel Dekker, Inc., New York, (1987) pp. 51-63].

5 The culture medium in which the hybridoma cells are cultured can then be assayed for the presence of monoclonal antibodies directed against PRO. Preferably, the binding specificity of monoclonal antibodies produced by the hybridoma cells is determined by immunoprecipitation or by an *in vitro* binding assay, such as radioimmunoassay (RIA) or enzyme-linked immunoabsorbent assay (ELISA). Such techniques and assays are known in the art. The binding affinity of the monoclonal antibody can, for example, be determined by the Scatchard analysis of Munson and Pollard, Anal. Biochem., 107:220 (1980).

10 After the desired hybridoma cells are identified, the clones may be subcloned by limiting dilution procedures and grown by standard methods [Goding, supra]. Suitable culture media for this purpose include, for example, Dulbecco's Modified Eagle's Medium and RPMI-1640 medium. Alternatively, the hybridoma cells may be grown *in vivo* as ascites in a mammal.

15 The monoclonal antibodies secreted by the subclones may be isolated or purified from the culture medium or ascites fluid by conventional immunoglobulin purification procedures such as, for example, protein A-Sepharose, hydroxylapatite chromatography, gel electrophoresis, dialysis, or affinity chromatography.

20 The monoclonal antibodies may also be made by recombinant DNA methods, such as those described in U.S. Patent No. 4,816,567. DNA encoding the monoclonal antibodies of the invention can be readily isolated and sequenced using conventional procedures (e.g., by using oligonucleotide probes that are capable of binding specifically to genes encoding the heavy and light chains of murine antibodies). The hybridoma cells of the invention serve as a preferred source of such DNA. Once isolated, the DNA may be placed into expression vectors, which are then transfected into host cells such as simian COS cells, Chinese hamster ovary (CHO) cells, or myeloma cells that do not otherwise produce immunoglobulin protein, to obtain the synthesis of monoclonal antibodies in the recombinant host cells. The DNA also may be modified, for example, by substituting the coding sequence for human heavy and light chain constant domains in place of the homologous murine sequences [U.S. Patent No. 4,816,567; Morrison et al., supra] or by covalently joining to the immunoglobulin coding sequence all or part of the coding sequence for a non-immunoglobulin polypeptide. Such a non-immunoglobulin polypeptide can be substituted for the constant domains of an antibody of the invention, or can be substituted for the variable domains of one antigen-combining site of an antibody of the invention to create a chimeric bivalent antibody.

30 The antibodies may be monovalent antibodies. Methods for preparing monovalent antibodies are well known in the art. For example, one method involves recombinant expression of immunoglobulin light chain and modified heavy chain. The heavy chain is truncated generally at any point in the Fc region so as to prevent heavy chain crosslinking. Alternatively, the relevant cysteine residues are substituted with another amino acid residue or are deleted so as to prevent crosslinking.

35 *In vitro* methods are also suitable for preparing monovalent antibodies. Digestion of antibodies to produce fragments thereof, particularly, Fab fragments, can be accomplished using routine techniques known in the art.

3. Human and Humanized Antibodies

The anti-PRO antibodies of the invention may further comprise humanized antibodies or human antibodies. Humanized forms of non-human (e.g., murine) antibodies are chimeric immunoglobulins, immunoglobulin chains or fragments thereof (such as Fv, Fab, Fab', F(ab')₂ or other antigen-binding subsequences of antibodies) which contain minimal sequence derived from non-human immunoglobulin.

5 Humanized antibodies include human immunoglobulins (recipient antibody) in which residues from a complementary determining region (CDR) of the recipient are replaced by residues from a CDR of a non-human species (donor antibody) such as mouse, rat or rabbit having the desired specificity, affinity and capacity. In some instances, Fv framework residues of the human immunoglobulin are replaced by corresponding non-human residues. Humanized antibodies may also comprise residues which are found neither in the recipient antibody

10 nor in the imported CDR or framework sequences. In general, the humanized antibody will comprise substantially all of at least one, and typically two, variable domains, in which all or substantially all of the CDR regions correspond to those of a non-human immunoglobulin and all or substantially all of the FR regions are those of a human immunoglobulin consensus sequence. The humanized antibody optimally also will comprise at least a portion of an immunoglobulin constant region (Fc), typically that of a human immunoglobulin [Jones

15 et al., Nature, 321:522-525 (1986); Riechmann et al., Nature, 332:323-329 (1988); and Presta, Curr. Op. Struct. Biol., 2:593-596 (1992)].

Methods for humanizing non-human antibodies are well known in the art. Generally, a humanized antibody has one or more amino acid residues introduced into it from a source which is non-human. These non-human amino acid residues are often referred to as "import" residues, which are typically taken from an "import" 20 variable domain. Humanization can be essentially performed following the method of Winter and co-workers [Jones et al., Nature, 321:522-525 (1986); Riechmann et al., Nature, 332:323-327 (1988); Verhoeyen et al., Science, 239:1534-1536 (1988)], by substituting rodent CDRs or CDR sequences for the corresponding sequences of a human antibody. Accordingly, such "humanized" antibodies are chimeric antibodies (U.S. Patent No. 4,816,567), wherein substantially less than an intact human variable domain has been substituted by the 25 corresponding sequence from a non-human species. In practice, humanized antibodies are typically human antibodies in which some CDR residues and possibly some FR residues are substituted by residues from analogous sites in rodent antibodies.

Human antibodies can also be produced using various techniques known in the art, including phage display libraries [Hoogenboom and Winter, J. Mol. Biol., 227:381 (1991); Marks et al., J. Mol. Biol., 222:581 30 (1991)]. The techniques of Cole et al. and Boerner et al. are also available for the preparation of human monoclonal antibodies (Cole et al., Monoclonal Antibodies and Cancer Therapy, Alan R. Liss, p. 77 (1985) and Boerner et al., J. Immunol., 147(1):86-95 (1991)]. Similarly, human antibodies can be made by introducing of human immunoglobulin loci into transgenic animals, e.g., mice in which the endogenous immunoglobulin genes have been partially or completely inactivated. Upon challenge, human antibody production is observed, 35 which closely resembles that seen in humans in all respects, including gene rearrangement, assembly, and antibody repertoire. This approach is described, for example, in U.S. Patent Nos. 5,545,807; 5,545,806; 5,569,825; 5,625,126; 5,633,425; 5,661,016, and in the following scientific publications: Marks et al.,

Bio/Technology 10, 779-783 (1992); Lonberg *et al.*, Nature 368 856-859 (1994); Morrison, Nature 368, 812-13 (1994); Fishwild *et al.*, Nature Biotechnology 14, 845-51 (1996); Neuberger, Nature Biotechnology 14, 826 (1996); Lonberg and Huszar, Intern. Rev. Immunol. 13 65-93 (1995).

4. Bispecific Antibodies

5 Bispecific antibodies are monoclonal, preferably human or humanized, antibodies that have binding specificities for at least two different antigens. In the present case, one of the binding specificities is for the PRO, the other one is for any other antigen, and preferably for a cell-surface protein or receptor or receptor subunit.

10 Methods for making bispecific antibodies are known in the art. Traditionally, the recombinant production of bispecific antibodies is based on the co-expression of two immunoglobulin heavy-chain/light-chain pairs, where the two heavy chains have different specificities [Milstein and Cuello, Nature, 305:537-539 (1983)]. Because of the random assortment of immunoglobulin heavy and light chains, these hybridomas (quadromas) produce a potential mixture of ten different antibody molecules, of which only one has the correct bispecific structure. The purification of the correct molecule is usually accomplished by affinity chromatography steps. 15 Similar procedures are disclosed in WO 93/08829, published 13 May 1993, and in Traunecker *et al.*, EMBO J., 10:3655-3659 (1991).

20 Antibody variable domains with the desired binding specificities (antibody-antigen combining sites) can be fused to immunoglobulin constant domain sequences. The fusion preferably is with an immunoglobulin heavy-chain constant domain, comprising at least part of the hinge, CH2, and CH3 regions. It is preferred to have the first heavy-chain constant region (CH1) containing the site necessary for light-chain binding present in at least one of the fusions. DNAs encoding the immunoglobulin heavy-chain fusions and, if desired, the immunoglobulin light chain, are inserted into separate expression vectors, and are co-transfected into a suitable host organism. For further details of generating bispecific antibodies see, for example, Suresh *et al.*, Methods in Enzymology, 121:210 (1986).

25 According to another approach described in WO 96/27011, the interface between a pair of antibody molecules can be engineered to maximize the percentage of heterodimers which are recovered from recombinant cell culture. The preferred interface comprises at least a part of the CH3 region of an antibody constant domain. In this method, one or more small amino acid side chains from the interface of the first antibody molecule are replaced with larger side chains (e.g. tyrosine or tryptophan). Compensatory "cavities" of identical or similar 30 size to the large side chain(s) are created on the interface of the second antibody molecule by replacing large amino acid side chains with smaller ones (e.g. alanine or threonine). This provides a mechanism for increasing the yield of the heterodimer over other unwanted end-products such as homodimers.

35 Bispecific antibodies can be prepared as full length antibodies or antibody fragments (e.g. F(ab')₂ bispecific antibodies). Techniques for generating bispecific antibodies from antibody fragments have been described in the literature. For example, bispecific antibodies can be prepared can be prepared using chemical linkage. Brennan *et al.*, Science 229:81 (1985) describe a procedure wherein intact antibodies are proteolytically cleaved to generate F(ab')₂ fragments. These fragments are reduced in the presence of the dithiol complexing

agent sodium arsenite to stabilize vicinal dithiols and prevent intermolecular disulfide formation. The Fab' fragments generated are then converted to thionitrobenzoate (TNB) derivatives. One of the Fab'-TNB derivatives is then reconverted to the Fab'-thiol by reduction with mercaptoethylamine and is mixed with an equimolar amount of the other Fab'-TNB derivative to form the bispecific antibody. The bispecific antibodies produced can be used as agents for the selective immobilization of enzymes.

5 Fab' fragments may be directly recovered from *E. coli* and chemically coupled to form bispecific antibodies. Shalaby *et al.*, J. Exp. Med. 175:217-225 (1992) describe the production of a fully humanized bispecific antibody F(ab')₂ molecule. Each Fab' fragment was separately secreted from *E. coli* and subjected to directed chemical coupling *in vitro* to form the bispecific antibody. The bispecific antibody thus formed was able to bind to cells overexpressing the ErbB2 receptor and normal human T cells, as well as trigger the lytic 10 activity of human cytotoxic lymphocytes against human breast tumor targets.

Various technique for making and isolating bispecific antibody fragments directly from recombinant cell culture have also been described. For example, bispecific antibodies have been produced using leucine zippers. Kostelnik *et al.*, J. Immunol. 148(5):1547-1553 (1992). The leucine zipper peptides from the Fos and Jun proteins were linked to the Fab' portions of two different antibodies by gene fusion. The antibody homodimers 15 were reduced at the hinge region to form monomers and then re-oxidized to form the antibody heterodimers. This method can also be utilized for the production of antibody homodimers. The "diabody" technology described by Hollinger *et al.*, Proc. Natl. Acad. Sci. USA 90:6444-6448 (1993) has provided an alternative mechanism for making bispecific antibody fragments. The fragments comprise a heavy-chain variable domain (V_H) connected to a light-chain variable domain (V_L) by a linker which is too short to allow pairing between the 20 two domains on the same chain. Accordingly, the V_H and V_L domains of one fragment are forced to pair with the complementary V_L and V_H domains of another fragment, thereby forming two antigen-binding sites. Another strategy for making bispecific antibody fragments by the use of single-chain Fv (sFv) dimers has also been reported. See, Gruber *et al.*, J. Immunol. 152:5368 (1994).

Antibodies with more than two valencies are contemplated. For example, trispecific antibodies can be prepared. 25 Tutt *et al.*, J. Immunol. 147:60 (1991).

Exemplary bispecific antibodies may bind to two different epitopes on a given PRO polypeptide herein. Alternatively, an anti-PRO polypeptide arm may be combined with an arm which binds to a triggering molecule on a leukocyte such as a T-cell receptor molecule (e.g. CD2, CD3, CD28, or B7), or Fc receptors for IgG (Fc γ R), such as Fc γ RI (CD64), Fc γ RII (CD32) and Fc γ RIII (CD16) so as to focus cellular defense mechanisms 30 to the cell expressing the particular PRO polypeptide. Bispecific antibodies may also be used to localize cytotoxic agents to cells which express a particular PRO polypeptide. These antibodies possess a PRO-binding arm and an arm which binds a cytotoxic agent or a radionuclide chelator, such as EOTUBE, DPTA, DOTA, or TETA. Another bispecific antibody of interest binds the PRO polypeptide and further binds tissue factor (TF).

35

5. Heteroconjugate Antibodies

Heteroconjugate antibodies are also within the scope of the present invention. Heteroconjugate

antibodies are composed of two covalently joined antibodies. Such antibodies have, for example, been proposed to target immune system cells to unwanted cells [U.S. Patent No. 4,676,980], and for treatment of HIV infection [WO 91/00360; WO 92/200373; EP 03089]. It is contemplated that the antibodies may be prepared *in vitro* using known methods in synthetic protein chemistry, including those involving crosslinking agents. For example, immunotoxins may be constructed using a disulfide exchange reaction or by forming a thioether bond.

5 Examples of suitable reagents for this purpose include iminothiolate and methyl-4-mercaptopbutyrimidate and those disclosed, for example, in U.S. Patent No. 4,676,980.

6. Effector Function Engineering

It may be desirable to modify the antibody of the invention with respect to effector function, so as to 10 enhance, *e.g.*, the effectiveness of the antibody in treating cancer. For example, cysteine residue(s) may be introduced into the Fc region, thereby allowing interchain disulfide bond formation in this region. The homodimeric antibody thus generated may have improved internalization capability and/or increased complement-mediated cell killing and antibody-dependent cellular cytotoxicity (ADCC). See Caron *et al.*, Exp Med., 176: 1191-1195 (1992) and Shopes, J. Immunol., 148: 2918-2922 (1992). Homodimeric antibodies 15 with enhanced anti-tumor activity may also be prepared using heterobifunctional cross-linkers as described in Wolff *et al.* Cancer Research, 53: 2560-2565 (1993). Alternatively, an antibody can be engineered that has dual Fc regions and may thereby have enhanced complement lysis and ADCC capabilities. See Stevenson *et al.*, Anti-Cancer Drug Design, 3: 219-230 (1989).

20 7. Immunoconjugates

The invention also pertains to immunoconjugates comprising an antibody conjugated to a cytotoxic agent such as a chemotherapeutic agent, toxin (*e.g.*, an enzymatically active toxin of bacterial, fungal, plant, or animal origin, or fragments thereof), or a radioactive isotope (*i.e.*, a radioconjugate).

Chemotherapeutic agents useful in the generation of such immunoconjugates have been described above.

25 Enzymatically active toxins and fragments thereof that can be used include diphtheria A chain, nonbinding active fragments of diphtheria toxin, exotoxin A chain (from *Pseudomonas aeruginosa*), ricin A chain, abrin A chain, modeccin A chain, alpha-sarcin, *Aleurites fordii* proteins, dianthin proteins, *Phytolaca americana* proteins (PAPI, PAPII, and PAP-S), momordica charantia inhibitor, curcin, crotin, sapaonaria officinalis inhibitor, gelonin, mitogellin, restrictocin, phenomycin, enomycin, and the trichothecenes. A variety of radionuclides are 30 available for the production of radioconjugated antibodies. Examples include ²¹²Bi, ¹³¹I, ¹³¹In, ⁹⁰Y, and ¹⁸⁶Re.

Conjugates of the antibody and cytotoxic agent are made using a variety of bifunctional protein-coupling agents such as N-succinimidyl-3-(2-pyridylthiol) propionate (SPDP), iminothiolane (IT), bifunctional derivatives of imidoesters (such as dimethyl adipimidate HCL), active esters (such as disuccinimidyl suberate), aldehydes (such as glutaraldehyde), bis-azido compounds (such as bis (p-azidobenzoyl) hexanediamine), bis-35 diazonium derivatives (such as bis-(p-diazoniumbenzoyl)-ethylenediamine), diisocyanates (such as tolyene 2,6-diisocyanate), and bis-active fluorine compounds (such as 1,5-difluoro-2,4-dinitrobenzene). For example, a ricin immunotoxin can be prepared as described in Vitetta *et al.*, Science, 238: 1098 (1987). Carbon-14-labeled 1-

isothiocyanatobenzyl-3-methyldiethylene triaminepentaacetic acid (MX-DTPA) is an exemplary chelating agent for conjugation of radionucleotide to the antibody. See WO94/11026.

In another embodiment, the antibody may be conjugated to a "receptor" (such streptavidin) for utilization in tumor pretargeting wherein the antibody-receptor conjugate is administered to the patient, followed by removal of unbound conjugate from the circulation using a clearing agent and then administration of a 5 "ligand" (e.g., avidin) that is conjugated to a cytotoxic agent (e.g., a radionucleotide).

8. Immunoliposomes

The antibodies disclosed herein may also be formulated as immunoliposomes. Liposomes containing the antibody are prepared by methods known in the art, such as described in Epstein *et al.*, Proc. Natl. Acad. Sci. USA, 82: 3688 (1985); Hwang *et al.*, Proc. Natl. Acad. Sci. USA, 77: 4030 (1980); and U.S. Pat. Nos. 10 4,485,045 and 4,544,545. Liposomes with enhanced circulation time are disclosed in U.S. Patent No. 4,013,556.

Particularly useful liposomes can be generated by the reverse-phase evaporation method with a lipid composition comprising phosphatidylcholine, cholesterol, and PEG-derivatized phosphatidylethanolamine(PEG-15 PE). Liposomes are extruded through filters of defined pore size to yield liposomes with the desired diameter. Fab' fragments of the antibody of the present invention can be conjugated to the liposomes as described in Martin *et al.*, J. Biol. Chem., 257: 286-288 (1982) via a disulfide-interchange reaction. A chemotherapeutic agent (such as Doxorubicin) is optionally contained within the liposome. See Gabizon *et al.*, J. National Cancer Inst., 81(19): 1484 (1989).

20

9. Pharmaceutical Compositions of Antibodies

Antibodies specifically binding a PRO polypeptide identified herein, as well as other molecules identified by the screening assays disclosed hereinbefore, can be administered for the treatment of various disorders in the form of pharmaceutical compositions.

25 If the PRO polypeptide is intracellular and whole antibodies are used as inhibitors, internalizing antibodies are preferred. However, lipofections or liposomes can also be used to deliver the antibody, or an antibody fragment, into cells. Where antibody fragments are used, the smallest inhibitory fragment that specifically binds to the binding domain of the target protein is preferred. For example, based upon the variable-region sequences of an antibody, peptide molecules can be designed that retain the ability to bind the target 30 protein sequence. Such peptides can be synthesized chemically and/or produced by recombinant DNA technology. See, *e.g.*, Marasco *et al.*, Proc. Natl. Acad. Sci. USA, 90: 7889-7893 (1993). The formulation herein may also contain more than one active compound as necessary for the particular indication being treated, preferably those with complementary activities that do not adversely affect each other. Alternatively, or in addition, the composition may comprise an agent that enhances its function, such as, for example, a cytotoxic 35 agent, cytokine, chemotherapeutic agent, or growth-inhibitory agent. Such molecules are suitably present in combination in amounts that are effective for the purpose intended.

The active ingredients may also be entrapped in microcapsules prepared, for example, by coacervation

techniques or by interfacial polymerization, for example, hydroxymethylcellulose or gelatin-microcapsules and poly-(methylmethacrylate) microcapsules, respectively, in colloidal drug delivery systems (for example, liposomes, albumin microspheres, microemulsions, nano-particles, and nanocapsules) or in macroemulsions. Such techniques are disclosed in Remington's *Pharmaceutical Sciences, supra*.

The formulations to be used for *in vivo* administration must be sterile. This is readily accomplished by 5 filtration through sterile filtration membranes.

Sustained-release preparations may be prepared. Suitable examples of sustained-release preparations include semipermeable matrices of solid hydrophobic polymers containing the antibody, which matrices are in the form of shaped articles, *e.g.*, films, or microcapsules. Examples of sustained-release matrices include polyesters, hydrogels (for example, poly(2-hydroxyethyl-methacrylate), or poly(vinylalcohol)), polylactides 10 (U.S. Pat. No. 3,773,919), copolymers of L-glutamic acid and γ ethyl-L-glutamate, non-degradable ethylene-vinyl acetate, degradable lactic acid-glycolic acid copolymers such as the LUPRON DEPOT TM (injectable microspheres composed of lactic acid-glycolic acid copolymer and leuprolide acetate), and poly-D(-)-3-hydroxybutyric acid. While polymers such as ethylene-vinyl acetate and lactic acid-glycolic acid enable release of molecules for over 100 days, certain hydrogels release proteins for shorter time periods. When encapsulated 15 antibodies remain in the body for a long time, they may denature or aggregate as a result of exposure to moisture at 37°C, resulting in a loss of biological activity and possible changes in immunogenicity. Rational strategies can be devised for stabilization depending on the mechanism involved. For example, if the aggregation mechanism is discovered to be intermolecular S-S bond formation through thio-disulfide interchange, stabilization may be achieved by modifying sulfhydryl residues, lyophilizing from acidic solutions, controlling moisture 20 content, using appropriate additives, and developing specific polymer matrix compositions.

G. Uses for anti-PRO Antibodies

The anti-PRO antibodies of the invention have various utilities. For example, anti-PRO antibodies may be used in diagnostic assays for PRO, *e.g.*, detecting its expression in specific cells, tissues, or serum. Various 25 diagnostic assay techniques known in the art may be used, such as competitive binding assays, direct or indirect sandwich assays and immunoprecipitation assays conducted in either heterogeneous or homogeneous phases [Zola, *Monoclonal Antibodies: A Manual of Techniques*, CRC Press, Inc. (1987) pp. 147-158]. The antibodies used in the diagnostic assays can be labeled with a detectable moiety. The detectable moiety should be capable of producing, either directly or indirectly, a detectable signal. For example, the detectable moiety may be a 30 radioisotope, such as 3 H, 14 C, 32 P, 35 S, or 125 I, a fluorescent or chemiluminescent compound, such as fluorescein isothiocyanate, rhodamine, or luciferin, or an enzyme, such as alkaline phosphatase, beta-galactosidase or horseradish peroxidase. Any method known in the art for conjugating the antibody to the detectable moiety may be employed, including those methods described by Hunter et al., *Nature*, 144:945 (1962); David et al., *Biochemistry*, 13:1014 (1974); Pain et al., *J. Immunol. Meth.*, 40:219 (1981); and Nygren, *J. Histochem. and 35 Cytochem.*, 30:407 (1982).

Anti-PRO antibodies also are useful for the affinity purification of PRO from recombinant cell culture or natural sources. In this process, the antibodies against PRO are immobilized on a suitable support, such a

Sephadex resin or filter paper, using methods well known in the art. The immobilized antibody then is contacted with a sample containing the PRO to be purified, and thereafter the support is washed with a suitable solvent that will remove substantially all the material in the sample except the PRO, which is bound to the immobilized antibody. Finally, the support is washed with another suitable solvent that will release the PRO from the antibody.

5 The following examples are offered for illustrative purposes only, and are not intended to limit the scope of the present invention in any way.

All patent and literature references cited in the present specification are hereby incorporated by reference in their entirety.

10

EXAMPLES

Commercially available reagents referred to in the examples were used according to manufacturer's instructions unless otherwise indicated. The source of those cells identified in the following examples, and throughout the specification, by ATCC accession numbers is the American Type Culture Collection, Rockville, Maryland.

15

EXAMPLE 1: Extracellular Domain Homology Screening to Identify Novel Polypeptides and cDNA Encoding Therefor

The extracellular domain (ECD) sequences (including the secretion signal sequence, if any) from about 950 known secreted proteins from the Swiss-Prot public database were used to search EST databases. The EST 20 databases included public databases (e.g., Dayhoff, GenBank), and proprietary databases (e.g. LIFESEQ™, Incyte Pharmaceuticals, Palo Alto, CA). The search was performed using the computer program BLAST or BLAST2 (Altschul, and Gish, Methods in Enzymology 266: 460-80 (1996); <http://blast.wustl.edu/blast/README.html>) as a comparison of the ECD protein sequences to a 6 frame translation of the EST sequences. Those comparisons with a Blast score of 70 (or in some cases 90) or greater 25 that did not encode known proteins were clustered and assembled into consensus DNA sequences with the program "phrap" (Phil Green, University of Washington, Seattle, Washington).

Using this extracellular domain homology screen, consensus DNA sequences were assembled relative 30 to the other identified EST sequences. In addition, the consensus DNA sequences obtained were often (but not always) extended using repeated cycles of BLAST and phrap to extend the consensus sequence as far as possible using the sources of EST sequences discussed above.

Based upon the consensus sequences obtained as described above, oligonucleotides were then synthesized and used to identify by PCR a cDNA library that contained the sequence of interest and for use as probes to isolate a clone of the full-length coding sequence for a PRO polypeptide. Forward (.f) and reverse (.r) PCR primers generally range from 20 to 30 nucleotides and are often designed to give a PCR product of 35 about 100-1000 bp in length. The probe (.p) sequences are typically 40-55 bp in length. In some cases, additional oligonucleotides are synthesized when the consensus sequence is greater than about 1-1.5kbp. In order to screen several libraries for a full-length clone, DNA from the libraries was screened by PCR amplification,

as per Ausubel et al., Current Protocols in Molecular Biology, with the PCR primer pair. A positive library was then used to isolate clones encoding the gene of interest using the probe oligonucleotide and one of the primer pairs.

The cDNA libraries used to isolate the cDNA clones were constructed by standard methods using commercially available reagents such as those from Invitrogen, San Diego, CA. The cDNA was primed with 5 oligo dT containing a NotI site, linked with blunt to Sall hemikinased adaptors, cleaved with NotI, sized appropriately by gel electrophoresis, and cloned in a defined orientation into a suitable cloning vector (such as pRKB or pRKD; pRKSb is a precursor of pRK5D that does not contain the SfiI site; see, Holmes et al., Science, 253:1278-1280 (1991)) in the unique Xhol and NotI sites.

10 **EXAMPLE 2: Isolation of cDNA Clones Encoding PRO211 and PRO217**

Consensus DNA sequences were assembled as described in Example 1 above and were designated as DNA28730 and DNA28760, respectively. Based on these consensus sequences, oligonucleotides were synthesized and used to identify by PCR a cDNA library that contained the sequences of interest and for use as probes to isolate a clone of the full-length coding sequence for the PRO211 and PRO217 polypeptides. The 15 libraries used to isolate DNA32292-1131 and DNA33094-1131 were fetal lung libraries.

cDNA clones were sequenced in their entirety. The entire nucleotide sequences of PRO211 (DNA32292-1131) and PRO217 (UNQ191) are shown in Figure 1 (SEQ ID NO: 1) and Figure 3 (SEQ ID NO:3), respectively. The predicted polypeptides are 353 and 379 amino acid in length, respectively, with respective molecular weights of approximately 38,190 and 41,520 daltons.

20 The oligonucleotide sequences used in the above procedures were the following:

28730.p (OLI 516) (SEQ ID NO:5)

5'-AGGGAGCACGGACAGTGTGCAGATGTGGACGAGTGCTCACTAGCA-3'

28730.f (OLI 517) (SEQ ID NO:6)

5'-AGAGTGTATCTCTGGCTACGC-3'

25 28730.r (OLI 518) (SEQ ID NO:7)

5'-TAAGTCCGGCACATTACAGGTC-3'

28760.p (OLI 617) (SEQ ID NO:8)

5'-CCCACGATGTATGAATGGTGGACTTGTGTGACTCCTGGTTCTGCATC-3'

28760.f (OLI 618) (SEQ ID NO:9)

30 5'-AAAGACGCATCTGCGAGTGTC-3'

28760.r (OLI 619) (SEQ ID NO:10)

5'-TGCTGATTCACACTGCTCTCCC-3'

EXAMPLE 3: Isolation of cDNA Clones Encoding Human PRO230

35 A consensus DNA sequence was assembled relative to the other identified EST sequences as described in Example 1 above, wherein the consensus sequence is designated herein as DNA30857. An EST proprietary to Genentech was employed in the consensus assembly. The EST is designated as DNA20088 and has the

nucleotide sequence shown in Figure 7 (SEQ ID NO:13).

Based on the DNA30857 consensus sequence, oligonucleotides were synthesized to identify by PCR a cDNA library that contained the sequence of interest and for use as probes to isolate a clone of the full-length coding sequence for PRO230.

A pair of PCR primers (forward and reverse) were synthesized:

5 forward PCR primer 5'-TTCGAGGCCTCTGAGAAGTGGCCC-3' (SEQ ID NO:14)
reverse PCR primer 5'-GGCGGTATCTCTGGCCTCCC-3' (SEQ ID NO:15)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA30857 sequence which had the following nucleotide sequence

hybridization probe

10 5'-TTCTCCACAGCAGCTGTGGCATCCGATCGTGTCAATCCATTCTCTGGG-3' (SEQ ID NO:16)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO230 gene using the probe oligonucleotide and one of the PCR primers.

15 RNA for construction of the cDNA libraries was isolated from human fetal lung tissue. DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO230 (herein designated as DNA33223-1136 and the derived protein sequence for PRO230.

The entire nucleotide sequence of DNA33223-1136 is shown in Figure 5 (SEQ ID NO:11). Clone DNA33223-1136 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 100-103 and ending at the stop codon at nucleotide positions 1501-1503 (Figure 5; SEQ ID NO:11).

20 The predicted polypeptide precursor is 467 amino acids long (Figure 6).

EXAMPLE 4: Isolation of cDNA Clones Encoding Human PRO232

A consensus DNA sequence was assembled relative to the other identified EST sequences as described in Example 1 above, wherein the consensus sequence is designated herein as DNA30935. Based on the 25 DNA30935 consensus sequence, oligonucleotides were synthesized to identify by PCR a cDNA library that contained the sequence of interest and for use as probes to isolate a clone of the full-length coding sequence for PRO232.

A pair of PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-TGCTGTGCTACTCCTGCAAAGCCC-3' (SEQ ID NO:19)
30 reverse PCR primer 5'-TGCACAAGTCGGTGTACAGCACG-3' (SEQ ID NO:20)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA30935 sequence which had the following nucleotide sequence

hybridization probe

5'-AGCAACGAGGACTGCCTGCAGGTGGAGAACTGCACCCAGCTGGG-3' (SEQ ID NO:21)

35 In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO232 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal kidney tissue.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO232 [herein designated as DNA34435-1140] and the derived protein sequence for PRO232.

The entire nucleotide sequence of DNA34435-1140 is shown in Figure 8 (SEQ ID NO:17). Clone DNA34435-1140 contains a single open reading frame with an apparent translational initiation site at nucleotide 5 positions 17-19 and ending at the stop codon at nucleotide positions 359-361 (Fig. 8; SEQ ID NO:17). The predicted polypeptide precursor is 114 amino acids long (Fig. 9). Clone DNA34435-1140 has been deposited with ATCC on September 16, 1997 and is assigned ATCC deposit no. ATCC 209250.

Analysis of the amino acid sequence of the full-length PRO232 suggests that it possesses 35% sequence identity with a stem cell surface antigen from *Gallus gallus*.

10

EXAMPLE 5: Isolation of cDNA Clones Encoding PRO187

A proprietary expressed sequence tag (EST) DNA database (LIFESEQ™, Incyte Pharmaceuticals, Palo Alto, CA) was searched and an EST (#843193) was identified which showed homology to fibroblast growth factor (FGF-8) also known as androgen-induced growth factor. mRNA was isolated from human fetal lung tissue 15 using reagents and protocols from Invitrogen, San Diego, CA (Fast Track 2). The cDNA libraries used to isolate the cDNA clones were constructed by standard methods using commercially available reagents (e.g., Invitrogen, San Diego, CA, Life Technologies, Gaithersburg, MD). The cDNA was primed with oligo dT containing a NotI site, linked with blunt to SalI hemikinased adaptors, cleaved with NotI, sized appropriately by gel electrophoresis, and cloned in a defined orientation into the cloning vector pRK5D using reagents and 20 protocols from Life Technologies, Gaithersburg, MD (Super Script Plasmid System). The double-stranded cDNA was sized to greater than 1000 bp and the SalI/NotI linker cDNA was cloned into XhoI/NotI cleaved vector. pRK5D is a cloning vector that has an sp6 transcription initiation site followed by an SfiI restriction enzyme site preceding the XhoI/NotI cDNA cloning sites.

Several libraries from various tissue sources were screened by PCR amplification with the following 25 oligonucleotide probes:

IN843193.f (OLI315) (SEQ ID NO:24)

5'-CAGTACGTGAGGGACCAGGGCGCCATGA-3'

IN843193.r (OLI 317) (SEQ ID NO:25)

5'-CCGGTGACCTGCACGTGCTTGCCA-3'

30 A positive library was then used to isolate clones encoding the PRO187 gene using one of the above oligonucleotides and the following oligonucleotide probe:

IN843193.p (OLI 316) (SEQ ID NO:26)

5'-GCGGATCTGCCGCCCTGCTCANCTGGTCGGTCATGGCGCCCT-3'

A cDNA clone was sequenced in entirety. The entire nucleotide sequence of PRO187 (DNA27864-35 1155) is shown in Figure 10 (SEQ ID NO:22). Clone DNA27864-1155 contains a single open reading frame with an apparent translational initiation site at nucleotide position 1 (Figure 10; SEQ ID NO:22). The predicted polypeptide precursor is 205 amino acids long. Clone DNA27864-1155 has been deposited with the ATCC

(designation: DNA27864-1155) and is assigned ATCC deposit no. ATCC 209375.

Based on a BLAST and FastA sequence alignment analysis (using the ALIGN computer program) of the full-length sequence, the PRO187 polypeptide shows 74% amino acid sequence identity (Blast score, 310) to human fibroblast growth factor-8 (androgen-induced growth factor).

5 **EXAMPLE 6: Isolation of cDNA Clones Encoding PRO265**

A consensus DNA sequence was assembled relative to other EST sequences as described in Example 1 above using phrap. This consensus sequence is herein designated DNA33679. Based on the DNA33679 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for

10 PRO265.

PCR primers (two forward and one reverse) were synthesized:

forward PCR primer A: 5'-CGGTCTACCTGTATGGCAACC-3' (SEQ ID NO:29);

forward PCR primer B: 5'-GCAGGACAACCAGATAAACAC-3' (SEQ ID NO:30);

reverse PCR primer 5'-ACGCAGATTGAGAAGGCTGTC-3' (SEQ ID NO:31)

15 Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA33679 sequence which had the following nucleotide sequence

hybridization probe

5'-TTCACGGGCTGCTCTGCCAGCTCTGAAGCTTGAAGAGCTGCAC-3' (SEQ ID NO:32)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was 20 screened by PCR amplification with PCR primer pairs identified above. A positive library was then used to isolate clones encoding the PRO265 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human a fetal brain library.

25 DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO265 [herein designated as DNA36350-1158] (SEQ ID NO:27) and the derived protein sequence for PRO265.

The entire nucleotide sequence of DNA36350-1158 is shown in Figure 12 (SEQ ID NO:27). Clone 30 DNA36350-1158 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 352-354 and ending at the stop codon at positions 2332-2334 (Figure 12). The predicted polypeptide precursor is 660 amino acids long (Figure 13). Clone DNA36350-1158 has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209378.

Analysis of the amino acid sequence of the full-length PRO265 polypeptide suggests that portions of it possess significant homology to the fibromodulin and the fibromodulin precursor, thereby indicating that PRO265 may be a novel member of the leucine rich repeat family, particularly related to fibromodulin.

35 **EXAMPLE 7: Isolation of cDNA Clones Encoding Human PRO219**

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is herein designated DNA28729. Based on the DNA28729

consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO219.

A pair of PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-GTGACCCTGGTTGTGAATACTCC-3' (SEQ ID NO:35)

5 reverse PCR primer 5'-ACAGCCATGGTCTATAGCTTGG-3' (SEQ ID NO:36)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA28729 sequence which had the following nucleotide sequence

hybridization probe

5'-GCCTGTCAGTGTCTGAGGGACACGTGCTCCGCAGCGATGGAAAG-3' (SEQ ID NO:37)

10 In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO219 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal kidney tissue.

15 DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO219 [herein designated as DNA32290-1164] (SEQ ID NO:33) and the derived protein sequence for PRO219.

20 The entire nucleotide sequence of DNA32290-1164 is shown in Figure 14 (SEQ ID NO:33). Clone DNA32290-1164 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 204-206 and ending at the stop codon at nucleotide positions 2949-2951 (Figure 14). The predicted polypeptide precursor is 915 amino acids long (Figure 15). Clone DNA32290-1164 has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209384.

Analysis of the amino acid sequence of the full-length PRO219 polypeptide suggests that portions of it possess significant homology to the mouse and human matrilin-2 precursor polypeptides.

25 EXAMPLE 8: Isolation of cDNA Clones Encoding Human PRO246

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is herein designated DNA30955. Based on the DNA30955 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO246.

30 A pair of PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-AGGGTCTCCAGGAGAAAGACTC-3' (SEQ ID NO:40)

reverse PCR primer 5'-ATTGTGGGCCTTGCAGACATAGAC-3' (SEQ ID NO:41)

35 Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA30955 sequence which had the following nucleotide sequence

hybridization probe

5'-GGCCACAGCATAAACCTTAGAACTCAATGTACTGGTCTCCAGCTCC-3' (SEQ ID NO:42)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO246 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal liver tissue. DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO246 [herein 5 designated as DNA35639-1172] (SEQ ID NO:38) and the derived protein sequence for PRO246.

The entire nucleotide sequence of DNA35639-1172 is shown in Figure 16 (SEQ ID NO:38). Clone DNA35639-1172 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 126-128 and ending at the stop codon at nucleotide positions 1296-1298 (Figure 16). The predicted polypeptide precursor is 390 amino acids long (Figure 17). Clone DNA35639-1172 has been deposited with 10 ATCC and is assigned ATCC deposit no. ATCC 209396.

Analysis of the amino acid sequence of the full-length PRO246 polypeptide suggests that it possess significant homology to the human cell surface protein HCAR, thereby indicating that PRO246 may be a novel cell surface virus receptor.

15 **EXAMPLE 9: Isolation of cDNA Clones Encoding Human PRO228**

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is herein designated DNA28758. An EST proprietary to Genentech was employed in the consensus assembly. This EST is shown in Figure 20 (SEQ ID NO:50) and is herein designated as DNA21951.

20 Based on the DNA28758 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO228.

PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-GGTAATGAGCTCCATTACAG-3' (SEQ ID NO:51)

25 forward PCR primer 5'-GGAGTAGAAAGCGCATGG-3' (SEQ ID NO:52)

forward PCR primer 5'-CACCTGATAACCATGAATGGCAG-3' (SEQ ID NO:53)

reverse PCR primer 5'-CGAGCTCGAATTAAATTCG-3' (SEQ ID NO:54)

reverse PCR primer 5'-GGATCTCCTGAGCTCAGG-3' (SEQ ID NO:55)

reverse PCR primer 5'-CCTAGTTGAGTGATCCTTGTAAG-3' (SEQ ID NO:56)

30 Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA28758 sequence which had the following nucleotide sequence

hybridization probe

5'-ATGAGACCCACACCTCATGCCGCTGTAATCACCTGACACATTGCAATT-3' (SEQ ID NO:57)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was 35 screened by PCR amplification with the PCR primer pairs identified above. A positive library was then used to isolate clones encoding the PRO228 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal kidney tissue.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO228 [herein designated as DNA33092-1202] (SEQ ID NO:48) and the derived protein sequence for PRO228.

5 The entire nucleotide sequence of DNA33092-1202 is shown in Figure 18 (SEQ ID NO:48). Clone DNA33092-1202 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 24-26 of SEQ ID NO:48 and ending at the stop codon after nucleotide position 2093 of SEQ ID NO:48. The predicted polypeptide precursor is 690 amino acids long (Figure 19). Clone DNA33092-1202 has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209420.

10 Analysis of the amino acid sequence of the full-length PRO228 polypeptide suggests that portions of it possess significant homology to the secretin-related proteins CD97 and EMR1 as well as the secretin member, latrophilin, thereby indicating that PRO228 may be a new member of the secretin related proteins.

EXAMPLE 10: Isolation of cDNA Clones Encoding Human PRO533

15 The EST sequence accession number AF007268, a murine fibroblast growth factor (FGF-15) was used to search various public EST databases (e.g., GenBank, Dayhoff, etc.). The search was performed using the computer program BLAST or BLAST2 [Altschul et al., *Methods in Enzymology*, 266:460-480 (1996); <http://blast.wustl.edu/blast/README.html>] as a comparison of the ECD protein sequences to a 6 frame translation of the EST sequences. The search resulted in a hit with GenBank EST AA220994, which has been identified as stratagene NT2 neuronal precursor 937230.

20 Based on the Genbank EST AA220994 sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence. Forward and reverse PCR primers may range from 20 to 30 nucleotides (typically about 24), and are designed to give a PCR product of 100-1000 bp in length. The probe sequences are typically 40-55 bp (typically about 50) in length. In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification, as per *Ausubel et al., Current Protocols in Molecular Biology*, with the PCR primer pair. A positive library was then used to isolate clones encoding the gene of interest using the probe oligonucleotide and one of the PCR primers.

25 In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified below. A positive library was then used to isolate clones encoding the PRO533 gene using the probe oligonucleotide and one of the PCR primers.

30 RNA for construction of the cDNA libraries was isolated from human fetal retina. The cDNA libraries used to isolated the cDNA clones were constructed by standard methods using commercially available reagents (e.g., Invitrogen, San Diego, CA; Clontech, etc.) The cDNA was primed with oligo dT containing a NotI site, linked with blunt to Sall hemikinased adaptors, cleaved with NotI, sized appropriately by gel electrophoresis, and cloned in a defined orientation into a suitable cloning vector (such as pRKB or pRKD; pRKB is a precursor 35 of pRK5D that does not contain the Sfil site; see, Holmes et al., *Science*, 253:1278-1280 (1991)) in the unique XhoI and NotI sites.

A cDNA clone was sequenced in its entirety. The full length nucleotide sequence of PRO533 is shown in Figure 21 (SEQ ID NO:58). Clone DNA49435-1219 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 459-461 (Figure 21; SEQ ID NO:58). The predicted polypeptide precursor is 216 amino acids long. Clone DNA47412-1219 has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209480.

5 Based on a BLAST-2 and FastA sequence alignment analysis of the full-length sequence, PRO533 shows amino acid sequence identity to fibroblast growth factor (53%).

The oligonucleotide sequences used in the above procedure were the following:

FGF15.forward: 5'-ATCCGCCAGATGGCTACAATGTGTA-3' (SEQ ID NO:60);

FGF15.probe: 5'-GCCTCCGGTCTCCCTGAGCAGTGCCAAACAGCGGCAGTGT-3' (SEQ ID NO:61);

10 FGF15.reverse: 5'-CCAGTCCGGTGACAAGCCAAA-3' (SEQ ID NO:62).

EXAMPLE 11: Isolation of cDNA Clones Encoding Human PRO245

A consensus DNA sequence was assembled relative to the other identified EST sequences as described in Example 1 above, wherein the consensus sequence is designated herein as DNA30954.

15 Based on the DNA30954 consensus sequence, oligonucleotides were synthesized to identify by PCR a cDNA library that contained the sequence of interest and for use as probes to isolate a clone of the full-length coding sequence for PRO245.

A pair of PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-ATCGTTGTGAAGTTAGTGCCCC-3' (SEQ ID NO:65)

20 reverse PCR primer 5'-ACCTGCGATATCCAACAGAACATTG-3' (SEQ ID NO:66)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA30954 sequence which had the following nucleotide sequence

hybridization probe

5'-GGAAGAGGATACAGTCACTCTGGAAGTATTAGTGGCTCCAGCAGTTCC-3' (SEQ ID NO:67)

25 In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO245 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal liver tissue. DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO245 [herein 30 designated as DNA35638-1141] and the derived protein sequence for PRO245.

The entire nucleotide sequence of DNA35638-1141 is shown in Figure 23 (SEQ ID NO:63). Clone DNA35638-1141 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 89-91 and ending at the stop codon at nucleotide positions 1025-1027 (Fig. 23; SEQ ID NO:63). The predicted polypeptide precursor is 312 amino acids long (Fig. 24). Clone DNA35638-1141 has been deposited 35 with ATCC on September 16, 1997 and is assigned ATCC deposit no. ATCC 209265.

Analysis of the amino acid sequence of the full-length PRO245 suggests that a portion of it possesses 60% amino acid identity with the human c-myb protein and, therefore, may be a new member of the

transmembrane protein receptor tyrosine kinase family.

EXAMPLE 12: Isolation of cDNA Clones Encoding Human PRO220, PRO221 and PRO227

(a) PRO220

A consensus DNA sequence was assembled relative to the other identified EST sequences as described

5 in Example 1 above, wherein the consensus sequence is designated herein as DNA28749. Based on the DNA28749 consensus sequence, oligonucleotides were synthesized to identify by PCR a cDNA library that contained the sequence of interest and for use as probes to isolate a clone of the full-length coding sequence for PRO220.

A pair of PCR primers (forward and reverse) were synthesized:

10 forward PCR primer 5'-TCACCTGGAGCCTTATTGCC-3' (SEQ ID NO:74)

reverse PCR primer 5'-ATACCAGCTATAACCAGGCTGCG-3' (SEQ ID NO:75)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA28749 sequence which had the following nucleotide sequence:

hybridization probe

15 5'-CAACAGTAAGTGGTTGATGCTCTTCAAATCTAGAGATTCTGATGATTGGG-3' (SEQ ID NO:76).

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO220 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal lung tissue. DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO220 (herein designated as DNA32298-1132 and the derived protein sequence for PRO220.

20 The entire nucleotide sequence of DNA32298-1132 is shown in Figure 25 (SEQ ID NO:68). Clone DNA32298-1132 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 480-482 and ending at the stop codon at nucleotide positions 2604-2606 (Figure 25). The predicted 25 polypeptide precursor is 708 amino acids long (Figure 26). Clone DNA32298-1132 has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209257.

Analysis of the amino acid sequence of the full-length PRO220 shows it has homology to member of the leucine rich repeat protein superfamily, including the leucine rich repeat protein and the neuronal leucine-rich repeat protein 1.

30

(b) PRO221

A consensus DNA sequence was assembled relative to the other identified EST sequences as described in Example 1 above, wherein the consensus sequence is designated herein as DNA28756. Based on the DNA28756 consensus sequence, oligonucleotides were synthesized to identify by PCR a cDNA library that contained the sequence of interest and for use as probes to isolate a clone of the full-length coding sequence for PRO221.

A pair of PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-CCATGTGTCTCCTCCTACAAAG-3' (SEQ ID NO:77)

reverse PCR primer 5'-GGGAATAGATGTGATCTGATTGG-3' (SEQ ID NO:78)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA28756 sequence which had the following nucleotide sequence:

hybridization probe

5 5'-CACCTGTAGCAAATCTCAAGGAAATACCTAGAGATCTTCCTCCTG-3' (SEQ ID NO:79)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO221 gene using the probe oligonucleotide and one of the PCR primers.

10 RNA for construction of the cDNA libraries was isolated from human fetal lung tissue. DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO221 [herein designated as DNA33089-1132 and the derived protein sequence for PRO221.

15 The entire nucleotide sequence of DNA33089-1132 is shown in Figure 27 (SEQ ID NO:70). Clone DNA33089-1132 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 179-181 and ending at the stop codon at nucleotide positions 956-958 (Figure 27). The predicted polypeptide precursor is 259 amino acids long (Figure 28). PRO221 is believed to have a transmembrane region at amino acids 206-225. Clone DNA33089-1132 has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209262.

Analysis of the amino acid sequence of the full-length PRO221 shows it has homology to member of the leucine rich repeat protein superfamily, including the SLIT protein.

20

(c) PRO227

A consensus DNA sequence was assembled relative to the other identified EST sequences as described in Example 1 above, wherein the consensus sequence is designated herein as DNA28740. Based on the DNA28740 consensus sequence, oligonucleotides were synthesized to identify by PCR a cDNA library that contained the sequence of interest and for use as probes to isolate a clone of the full-length coding sequence for PRO227.

A pair of PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-AGCAACCGCCTGAAGCTCATCC-3' (SEQ ID NO:80)

reverse PCR primer 5'-AAGGCGCGGTGAAAGATGTAGACG-3' (SEQ ID NO:81)

30 Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA28740 sequence which had the following nucleotide sequence:

hybridization probe

5'GACTACATGTTCAAGGACCTGTACAACCTCAAGTCACTGGAGGTTGGCGA-3' (SEQ ID NO:82).

35 In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO227 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal lung tissue. DNA

sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO227 [herein designated as DNA33786-1132 and the derived protein sequence for PRO227.

The entire nucleotide sequence of DNA33786-1132 is shown in Figure 29 (SEQ ID NO:72). ,Clone DNA33786-1132 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 33-35 and ending at the stop codon at nucleotide positions 1893-1895 (Figure 29). The predicted 5 polypeptide precursor is 620 amino acids long (Figure 30). PRO227 is believed to have a transmembrane region. Clone DNA33786-1132 has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209253.

Analysis of the amino acid sequence of the full-length PRO221 shows it has homology to member of the leucine rich repeat protein superfamily, including the platelet glycoprotein V precursor and the human 10 glycoprotein V.

EXAMPLE 13: Isolation of cDNA Clones Encoding Human PRO258

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is herein designated DNA28746.

15 Based on the DNA28746 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO258.

PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-GCTAGGAATTCCACAGAACGCC-3' (SEQ ID NO:85)

20 reverse PCR primer 5'-AACCTGGAATGTCACCGAGCTG-3' (SEQ ID NO:86)

reverse PCR primer 5'-CCTAGCACAGTGACGAGGGACTTGGC-3' (SEQ ID NO:87)

Additionally, synthetic oligonucleotide hybridization probes were constructed from the consensus DNA28740 sequence which had the following nucleotide sequence:

hybridization probe

25 5'-AAGACACAGCCACCCCTAAACTGTCAGTCTTCTGGGAGCAAGCCTGCAGCC-3' (SEQ ID NO:88)

5'-GCCCTGGCAGACGAGGGCGAGTACACCTGCTCAATCTCACTATGCCTGT-3' (SEQ ID NO:89)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO258 gene using the probe oligonucleotide and one of the PCR primers.

30 RNA for construction of the cDNA libraries was isolated from human fetal lung tissue. DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO258 [herein designated as DNA35918-1174] (SEQ ID NO:83) and the derived protein sequence for PRO258.

The entire nucleotide sequence of DNA35918-1174 is shown in Figure 31 (SEQ ID NO:83). Clone DNA35918-1174 contains a single open reading frame with an apparent translational initiation site at nucleotide 35 positions 147-149 of SEQ ID NO:83 and ending at the stop codon after nucleotide position 1340 of SEQ ID NO:83 (Figure 31). The predicted polypeptide precursor is 398 amino acids long (Figure 32). Clone DNA35918-1174 has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209402.

Analysis of the amino acid sequence of the full-length PRO258 polypeptide suggests that portions of it possess significant homology to the CRTAM and the poliovirus receptor and have an Ig domain, thereby indicating that PRO258 is a new member of the Ig superfamily.

EXAMPLE 14: Isolation of cDNA Clones Encoding Human PRO266

5 An expressed sequence tag database was searched for ESTs having homology to SLIT, resulting in the identification of a single EST sequence designated herein as T73996. Based on the T73996 EST sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO266.

A pair of PCR primers (forward and reverse) were synthesized:

10 forward PCR primer 5'-GTTGGATCTGGCAACAATAAC-3' (SEQ ID NO:92)

reverse PCR primer 5'-ATTGTTGTGCAGGCTGAGTTAAG-3' (SEQ ID NO:93)

Additionally, a synthetic oligonucleotide hybridization probe was constructed which had the following nucleotide sequence

hybridization probe

15 5'-GGTGGCTATACATGGATAGCAATTACCTGGACACGCTGTCCCGGG-3' (SEQ ID NO:94)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO266 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal brain tissue. DNA 20 sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO266 [herein designated as DNA37150-1178] (SEQ ID NO:90) and the derived protein sequence for PRO266..

The entire nucleotide sequence of DNA37150-1178 is shown in Figure 33 (SEQ ID NO:90). Clone DNA37150-1178 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 167-169 and ending at the stop codon after nucleotide position 2254 of SEQ ID NO:90. The predicted 25 polypeptide precursor is 696 amino acids long (Figure 34). Clone DNA37150-1178 has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209401.

Analysis of the amino acid sequence of the full-length PRO266 polypeptide suggests that portions of it possess significant homology to the SLIT protein, thereby indicating that PRO266 may be a novel leucine rich repeat protein.

30

EXAMPLE 15: Isolation of cDNA Clones Encoding Human PRO269

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is herein designated DNA35705. Based on the DNA35705 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO269.

Forward and reverse PCR primers were synthesized:

forward PCR primer (.f1) 5'-TGGAAAGGAGATGCGATGCCACCTG -3'
(SEQ ID NO:97)

forward PCR primer (.f2) 5'-TGACCAGTGGGAAGGACAG-3' (SEQ ID NO:98)

forward PCR primer (.f3) 5'-ACAGAGCAGAGGGTGCCTG-3' (SEQ ID NO:99)

reverse PCR primer (r1) 5'-TCAGGGACAAGTGGTGTCTCTCCC-3'

5 (SEQ ID NO:100)

reverse PCR primer (r2) 5'-TCAGGGAAAGGAGTGTGCAGTTCTG-3'
(SEQ ID NO:101)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA35705 sequence which had the following nucleotide sequence:

10 hybridization probe
5'-ACAGCTCCGATCTCAGTTACTTGCATCGCGACGAAATCGCGCTCGCT-3' (SEQ ID NO:102)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pairs identified above. A positive library was then used to isolate clones encoding the PRO269 gene using the probe oligonucleotide and one of the PCR primers.

15 RNA for construction of the cDNA libraries was isolated from human fetal kidney tissue.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO269 [herein designated as DNA38260-1180] (SEQ ID NO:95) and the derived protein sequence for PRO269.

The entire nucleotide sequence of DNA38260-1180 is shown in Figure 35 (SEQ ID NO:95). Clone 20 DNA38260-1180 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 314-316 and ending at the stop codon at nucleotide positions 1784-1786 (Fig. 35; SEQ ID NO:95). The predicted polypeptide precursor is 490 amino acids long (Fig. 36). Clone DNA38260-1180 has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209397.

Analysis of the amino acid sequence of the full-length PRO269 suggests that portions of it possess 25 significant homology to the human thrombomodulin proteins, thereby indicating that PRO269 may possess one or more thrombomodulin-like domains.

EXAMPLE 16: Isolation of cDNA Clones Encoding Human PRO287

A consensus DNA sequence encoding PRO287 was assembled relative to the other identified EST 30 sequences as described in Example 1 above, wherein the consensus sequence is designated herein as DNA28728. Based on the DNA28728 consensus sequence, oligonucleotides were synthesized to identify by PCR a cDNA library that contained the sequence of interest and for use as probes to isolate a clone of the full-length coding sequence for PRO287.

A pair of PCR primers (forward and reverse) were synthesized:

35 forward PCR primer 5'-CCGATTAGACCTCGAGAGT-3' (SEQ ID NO:105)
reverse PCR primer 5'-GTCAAGGAGTCCTCCACAATAC-3' (SEQ ID NO:106)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA28728

sequence which had the following nucleotide sequence

hybridization probe

5' -GTGTACAATGGCCATGCCAATGGCCAGCGCATTGGCCGCTTCTGT-3'

(SEQ ID NO:107)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was
5 screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to
isolate clones encoding the PRO287 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal kidney tissue.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for
10 PRO287 [herein designated as DNA39969-1185, SEQ ID NO:103] and the derived protein sequence for
PRO287.

The entire nucleotide sequence of DNA39969-1185 is shown in Figure 37 (SEQ ID NO:103). Clone
DNA39969-1185 contains a single open reading frame with an apparent translational initiation site at nucleotide
positions 307-309 and ending at the stop codon at nucleotide positions 1552-1554 (Fig. 37; SEQ ID NO:103).
The predicted polypeptide precursor is 415 amino acids long (Fig. 38). Clone DNA39969-1185 has been
15 deposited with ATCC and is assigned ATCC deposit no. ATCC 209400.

Analysis of the amino acid sequence of the full-length PRO287 suggests that it may possess one or more
procollagen C-proteinase enhancer protein precursor or procollagen C-proteinase enhancer protein-like domains.
Based on a BLAST and FastA sequence alignment analysis of the full-length sequence, PRO287 shows nucleic
acid sequence identity to procollagen C-proteinase enhancer protein precursor and procollagen C-proteinase
20 enhancer protein (47 and 54%, respectively).

EXAMPLE 17: Isolation of cDNA Clones Encoding Human PRO214

A consensus DNA sequence was assembled using phrap as described in Example 1 above. This
consensus DNA sequence is designated herein as DNA28744. Based on this consensus sequence,
25 oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest,
and 2) for use as probes to isolate a clone of the full-length coding sequence.

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was
screened by PCR amplification with the PCR primer pair identified below. A positive library was then used to
isolate clones encoding the PRO214 gene using the probe oligonucleotide and one of the PCR primers.

30 RNA for construction of the cDNA libraries was isolated from human fetal lung tissue.

A cDNA clone was sequenced in its entirety. The full length nucleotide sequence of DNA32286-1191
is shown in Figure 39 (SEQ ID NO:108). DNA32286-1191 contains a single open reading frame with an
apparent translational initiation site at nucleotide position 103 (Fig. 39; SEQ ID NO:108). The predicted
polypeptide precursor is 420 amino acids long (SEQ ID NO:109).

35 Based on a BLAST and FastA sequence alignment analysis of the full-length sequence, PRO214
polypeptide shows amino acid sequence identity to HT protein and/or Fibulin (49% and 38%, respectively).

The oligonucleotide sequences used in the above procedure were the following:

28744.p (OLI555)
5'-CCTGGCTATCAGCAGGTGGGCTCCAAGTGTCTCGATGTGGATGAGTGTGA-3' (SEQ ID NO:110)

28744.f (OLI556)
5'-ATTCTGCGTGAACACTGAGGGC-3' (SEQ ID NO:111)

28744.r (OLI557)
5 5'-ATCTGCTTGTAGCCCTCGGCAC-3' (SEQ ID NO:112)

EXAMPLE 18: Isolation of cDNA Clones Encoding Human PRO317

A consensus DNA sequence was assembled using phrap as described in Example 1 above, wherein the consensus sequence is herein designated as DNA28722. Based on this consensus sequence, oligonucleotides 10 were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence. The forward and reverse PCR primers, respectively, synthesized for this purpose were:

5'-AGGACTGCCATAACTGCCCTG (OLI489) (SEQ ID NO:115) and
5'-ATAGGAGTTGAAGCAGCGCTGC (OLI490) (SEQ ID NO:116).

15 The probe synthesized for this purpose was:

5'-TGTGTGGACATAGACGAGTGCCGCTACCGCTACTGCCAGCACCGC (OLI488) (SEQ ID NO:117)

mRNA for construction of the cDNA libraries was isolated from human fetal kidney tissue.

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was 20 screened by PCR amplification, as per Ausubel *et al.*, *Current Protocols in Molecular Biology* (1989), with the PCR primer pair identified above. A positive library was then used to isolate clones containing the PRO317 gene using the probe oligonucleotide identified above and one of the PCR primers.

A cDNA clone was sequenced in its entirety. The entire nucleotide sequence of DNA33461-1199 (encoding PRO317) is shown in Figure 41 (SEQ ID NO:113). Clone DNA33461-1199 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 68-70 (Fig. 41; SEQ ID 25 NO:113). The predicted polypeptide precursor is 366 amino acids long. The predicted signal sequence is amino acids 1-18 of Figure 42 (SEQ ID NO:114). There is one predicted N-linked glycosylation site at amino acid residue 160. Clone DNA33461-1199 has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209367.

Based on BLAST™ and FastA™ sequence alignment analysis (using the ALIGN™ computer program) 30 of the full-length PRO317 sequence, PRO317 shows the most amino acid sequence identity to EBAF-1 (92%). The results also demonstrate a significant homology between human PRO317 and mouse LEFTY protein. The C-terminal end of the PRO317 protein contains many conserved sequences consistent with the pattern expected of a member of the TGF- superfamily.

In situ expression analysis in human tissues performed as described below evidences that there is 35 distinctly strong expression of the PRO317 polypeptide in pancreatic tissue.

EXAMPLE 19: Isolation of cDNA clones Encoding Human PRO301

A consensus DNA sequence designated herein as DNA35936 was assembled using phrap as described in Example 1 above. Based on this consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence.

5 In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified below. A positive library was then used to isolate clones encoding the PRO301 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal kidney.

10 A cDNA clone was sequenced in its entirety. The full length nucleotide sequence of native sequence PRO301 is shown in Figure 43 (SEQ ID NO:118). Clone DNA40628-1216 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 52-54 (Fig. 43; SEQ ID NO:118). The predicted polypeptide precursor is 299 amino acids long with a predicted molecular weight of 32,583 daltons and pI of 8.29. Clone DNA40628-1216 has been deposited with ATCC and is assigned ATCC deposit No. ATCC 209432.

15 Based on a BLAST and FastA sequence alignment analysis of the full-length sequence, PRO301 shows amino acid sequence identity to A33 antigen precursor (30%) and coxsackie and adenovirus receptor protein (29%).

The oligonucleotide sequences used in the above procedure were the following:

20 OLI2162 (35936.f1) 5'-TCGCGGAGCTGTGTTCTGTTCCC-3' (SEQ ID NO:120)
OLI2163 (35936.p1)
5'-TGATCGCGATGGGGACAAAGGCAGCTCGAGAGGAACTGTTGTCCT-3' (SEQ ID NO:121)
OLI2164 (35936.f2)
5'-ACACCTGGTCAAAGATGGG-3' (SEQ ID NO:122)
OLI2165 (35936.r1)
25 5'-TAGGAAGAGTTGCTGAAGGCACGG-3' (SEQ ID NO:123)
OLI2166 (35936.f3)
5'-TTGCCTTACTCAGGTGCTAC-3' (SEQ ID NO:124)
OLI2167 (35936.r2)
5'-ACTCAGCAGTGGTAGGAAAG-3' (SEQ ID NO:125)
30

EXAMPLE 20: Isolation of cDNA Clones Encoding Human PRO224

A consensus DNA sequence assembled relative to the other identified EST sequences as described in Example 1, wherein the consensus sequence is designated herein as DNA30845. Based on the DNA30845 consensus sequence, oligonucleotides were synthesized to identify by PCR a cDNA library that contained the sequence of interest and for use as probes to isolate a clone of the full-length coding sequence for PRO224.

35 A pair of PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-AAGTTCCAGTGCGCACCGAGTGGC-3' (SEQ ID NO:128)
reverse PCR primer 5'-TTGGTTCCACAGCCGAGCTCGTCG-3' (SEQ ID NO:129)

40 Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA30845 sequence which had the following nucleotide sequence
hybridization probe

5'-GAGGAGGTGCAGGATTGAGCCATGTACCCAGAAAGGGCAATGCCACC-3' (SEQ ID NO:130)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO224 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal liver tissue.

5 DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO224 [herein designated as DNA33221-1133] and the derived protein sequence for PRO224.

10 The entire nucleotide sequence of DNA33221-1133 is shown in Figure 45 (SEQ ID NO:126). Clone DNA33221-1133 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 33-35 and ending at the stop codon at nucleotide positions 879-899 (Figure 45; SEQ ID NO:126). The start of a transmembrane region begins at nucleotide position 777. The predicted polypeptide precursor is 282 amino acids long (Figure 46). Clone DNA33221-1133 has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209263.

15 Analysis of the amino acid sequence of the full-length PRO224 suggests that it has homology to very low-density lipoprotein receptors, apolipoprotein E receptor and chicken oocyte receptors P95. Based on a BLAST and FastA sequence alignment analysis of the full-length sequence, PRO224 has amino acid identity to portions of these proteins in the range from 28% to 45%, and overall identity with these proteins in the range from 33% to 39%.

EXAMPLE 21: Isolation of cDNA Clones Encoding Human PRO222

20 A consensus DNA sequence was assembled relative to the other identified EST sequences as described in Example 1 above, wherein the consensus sequence is designated herein as DNA28771. Based on the DNA28771 consensus sequence, oligonucleotides were synthesized to identify by PCR a cDNA library that contained the sequence of interest and for use as probes to isolate a clone of the full-length coding sequence for PRO222.

25 A pair of PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-ATCTCCTATCGCTGCTTCCGG-3' (SEQ ID NO:133)

reverse PCR primer 5'-AGCCAGGATCGCAGTAAACTCC-3' (SEQ ID NO:134)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA28771 sequence which had the following nucleotide sequence:

30 hybridization probe

5'-ATTTAAACTTGATGGGTCTGCGTATCTTGAGTGCTTACAAAACCTTATCT-3' (SEQ ID NO:135)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO222 gene using the probe oligonucleotide and one of the PCR primers.

35 RNA for construction of the cDNA libraries was isolated from human fetal kidney tissue.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO222 [herein designated as DNA33107-1135] and the derived protein sequence for PRO222.

The entire nucleotide sequence of DNA33107-1135 is shown in Figure 47 (SEQ ID NO:131). Clone DNA33107-1135 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 159-161 and ending at the stop codon at nucleotide positions 1629-1631 (Fig. 47; SEQ ID NO:131). The predicted polypeptide precursor is 490 amino acids long (Fig. 48). Clone DNA33107-1135 has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209251.

5 Based on a BLAST and FastA sequence alignment analysis of the full-length sequence, PRO222 shows amino acid sequence identity to mouse complement factor h precursor (25-26%), complement receptor (27-29%), mouse complement C3b receptor type 2 long form precursor (25-47%) and human hypothetical protein kiaa0247 (40%).

10 **EXAMPLE 22: Isolation of cDNA clones Encoding PRO234**

A consensus DNA sequence was assembled (DNA30926) using phrap as described in Example 1 above. Based on this consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence.

15 RNA for the construction of the cDNA libraries was isolated using standard isolation protocols, e.g., Ausubel *et al.*, *Current Protocols in Molecular Biology*, from tissue or cell line sources or it was purchased from commercial sources (e.g., Clontech). The cDNA libraries used to isolate the cDNA clones were constructed by standard methods (e.g., Ausubel *et al.*) using commercially available reagents (e.g., Invitrogen). This library was derived from 22 week old fetal brain tissue.

20 A cDNA clone was sequenced in its entirety. The entire nucleotide sequence of PRO234 is shown in Figure 49 (SEQ ID NO:136). The predicted polypeptide precursor is 382 amino acids long and has a calculated molecular weight of approximately 43.1 kDa.

The oligonucleotide sequences used in the above procedure were the following:

30926.p (OLI826) (SEQ ID NO:138): 5'-GTTCAATTGAAAACCTCTTGCCATCT
GATGGTGACTTCTGGATTGGGCTCA-3'

25 30926.f (OLI827) (SEQ ID NO:139): 5'-AAGCCAAAGAAGCCTGCAGGAGGG-3'
30926.r (OLI828) (SEQ ID NO:140): 5'-CAGTCCAAGCATAAAGGTCTGGC-3'

EXAMPLE 23: Isolation of cDNA Clones Encoding Human PRO231

A consensus DNA sequence was assembled relative to the other identified EST sequences as described in Example 1 above, wherein the consensus sequence was designated herein as DNA30933. Based on the DNA30933 consensus sequence, oligonucleotides were synthesized to identify by PCR a cDNA library that contained the sequence of interest and for use as probes to isolate a clone of the full-length coding sequence for PRO231.

Three PCR primers (two forward and one reverse) were synthesized:

35 forward PCR primer 1 5'-CCAATACCAAAAGCTGGAGCC-3' (SEQ ID NO:143)
forward PCR primer 2 5'-GCAGCTCTATTACCACGGGAAGGA-3' (SEQ ID NO:144)
reverse PCR primer 5'-TCCTTCCCGTGGTAATAGAGCTGC-3' (SEQ ID NO:145)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA30933 sequence which had the following nucleotide sequence

hybridization probe

5' -GGCAGAGAACCAAGAGGCCGGAGGAGACTGCCTCTTACAGCCAGG-3' (SEQ ID NO:146)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was
5 screened by PCR amplification with the PCR primer pairs identified above. A positive library was then used to isolate clones encoding the PRO231 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal liver tissue.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO231 [herein designated as DNA34434-1139] and the derived protein sequence for PRO231.

10 The entire nucleotide sequence of DNA34434-1139 is shown in Figure 51 (SEQ ID NO:141). Clone DNA34434-1139 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 173-175 and ending at the stop codon at nucleotide positions 1457-1459 (Fig. 51; SEQ ID NO:141). The predicted polypeptide precursor is 428 amino acids long (Fig. 52). Clone DNA34434-1139 has been deposited with ATCC on September 16, 1997 and is assigned ATCC deposit no. ATCC 209252.

15 Analysis of the amino acid sequence of the full-length PRO231 suggests that it possesses 30% and 31% amino acid identity with the human and rat prostatic acid phosphatase precursor proteins, respectively.

EXAMPLE 24: Isolation of cDNA Clones Encoding Human PRO229

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described
20 in Example 1 above. This consensus sequence is herein designated DNA28762. Based on the DNA28762 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO229.

A pair of PCR primers (forward and reverse) were synthesized:

25 forward PCR primer 5' -TTCAGCTCATCACCTTCACCTGCC-3' (SEQ ID NO:149)

reverse PCR primer 5' -GGCTCATACAAATACCACTAGGG-3' (SEQ ID NO:150)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA28762 sequence which had the following nucleotide sequence

hybridization probe

30 5' -GGGCCTCCACCGCTGTGAAGGGCGGGTGGAGGTGGAACAGAAAGGCCAGT-3' (SEQ ID NO:151)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO229 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal liver tissue.

35 DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO229 [herein designated as DNA33100-1159] (SEQ ID NO:147) and the derived protein sequence for PRO229.

The entire nucleotide sequence of DNA33100-1159 is shown in Figure 53 (SEQ ID NO:147). Clone DNA33100-1159 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 98-100 and ending at the stop codon at nucleotide positions 1139-1141 (Figure 53). The predicted polypeptide precursor is 347 amino acids long (Figure 54). Clone DNA33100-1159 has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209377

5 Analysis of the amino acid sequence of the full-length PRO229 polypeptide suggests that portions of it possess significant homology to antigen wc1.1, M130 antigen and CD6.

EXAMPLE 25: Isolation of cDNA Clones Encoding Human PRO238

10 A consensus DNA sequence was assembled relative to other EST sequences using phrap as described above in Example 1. This consensus sequence is herein designated DNA30908. Based on the DNA30908 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO238.

PCR primers (forward and reverse) were synthesized:

15 forward PCR primer 1 5'-GGTGCTAAACTGGTGCTCTGTGGC-3' (SEQ ID NO:154)

forward PCR primer 2 5'-CAGGGCAAGATGAGCATTCC-3' (SEQ ID NO:155)

reverse PCR primer 5'-TCATACTGTTCCATCTCGGCACGC-3' (SEQ ID NO:156)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA30908 sequence which had the following nucleotide sequence

20 hybridization probe

5'-AATGGTGGGCCCTAGAAGAGCTCATCAGAGAACTCACCGCTTCTCATGC-3' (SEQ ID NO:157)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO238 gene using the probe oligonucleotide and one of the PCR primers.

25 RNA for construction of the cDNA libraries was isolated from human fetal liver tissue.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO238 and the derived protein sequence for PRO238.

30 The entire nucleotide sequence of DNA35600-1162 is shown in Figure 55 (SEQ ID NO:152). Clone DNA35600-1162 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 134-136 and ending prior to the stop codon at nucleotide positions 1064-1066 (Figure 55). The predicted polypeptide precursor is 310 amino acids long (Figure 56). Clone DNA35600-1162 has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209370.

35 Analysis of the amino acid sequence of the full-length PRO238 polypeptide suggests that portions of it possess significant homology to reductase, particularly oxidoreductase, thereby indicating that PRO238 may be a novel reductase.

EXAMPLE 26: Isolation of cDNA Clones Encoding Human PRO233

The extracellular domain (ECD) sequences (including the secretion signal, if any) of from about 950 known secreted proteins from the Swiss-Prot public protein database were used to search expressed sequence tag (EST) databases. The EST databases included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ™, Incyte Pharmaceuticals, Palo Alto, CA). The search was performed using the 5 computer program BLAST or BLAST2 (Altshul et al., *Methods in Enzymology* 266:460-480 (1996)) as a comparison of the ECD protein sequences to a 6 frame translation of the EST sequence. Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into consensus DNA sequences with the program "phrap" (Phil Green, University of Washington, Seattle, Washington; <http://bozeman.mbt.washington.edu/phrap.docs/phrap.html>).

10 An expressed sequence tag (EST) was identified by the EST database search and a consensus DNA sequence was assembled relative to other EST sequences using phrap. This consensus sequence is herein designated DNA30945. Based on the DNA30945 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO233.

15 Forward and reverse PCR primers were synthesized:

forward PCR primer 5'-GGTGAAGGCAGAAATTGGAGATG-3' (SEQ ID NO:160)

reverse PCR primer 5'-ATCCCATGCATCAGCCTGTTACC-3' (SEQ ID NO:161)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA30945 sequence which had the following nucleotide sequence

20 hybridization probe

5'-GCTGGTGTAGTCTATACATCAGATTGTTGCTACACAAGATCCTCAG-3'

(SEQ ID NO:162)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to 25 isolate clones encoding the PRO233 gene using the probe oligonucleotide.

RNA for construction of the cDNA libraries was isolated from human fetal brain tissue.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO233 [herein designated as DNA34436-1238] (SEQ ID NO:158) and the derived protein sequence for PRO233.

30 The entire nucleotide sequence of DNA34436-1238 is shown in Figure 57 (SEQ ID NO:158). Clone DNA34436-1238 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 101-103 and ending at the stop codon at nucleotide positions 1001-1003 (Figure 57). The predicted polypeptide precursor is 300 amino acids long (Figure 58). The full-length PRO233 protein shown in Figure 58 has an estimated molecular weight of about 32,964 daltons and a pI of about 9.52. Clone DNA34436-1238 35 has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209523.

Analysis of the amino acid sequence of the full-length PRO233 polypeptide suggests that portions of it possess significant homology to reductase proteins, thereby indicating that PRO233 may be a novel reductase.

EXAMPLE 27: Isolation of cDNA Clones Encoding Human PRO223

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is herein designated DNA30836. Based on the DNA30836 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO223.

5 PCR primer pairs (one forward and two reverse) were synthesized:

forward PCR primer 5'-TTCCATGCCACCTAAGGGAGACTC-3' (SEQ ID NO:165)

reverse PCR primer 1 5'-TGGATGAGGTGTGCAATGGCTGGC-3' (SEQ ID NO:166)

reverse PCR primer 2 5'-AGCTCTCAGAGGCTGGTCATAGGG-3' (SEQ ID NO:167)

10 Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA30836 sequence which had the following nucleotide sequence

hybridization probe

5'-GTCGGCCCTTCCCAGGACTGAACATGAAGAGTTATGCCGGCTTCCTCAC-3' (SEQ ID NO:168)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was 15 screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO223 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal liver tissue.

20 DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO223 [herein designated as DNA33206-1165] (SEQ ID NO:163) and the derived protein sequence for PRO223.

The entire nucleotide sequence of DNA33206-1165 is shown in Figure 59 (SEQ ID NO:163). Clone DNA33206-1165 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 97-99 and ending at the stop codon at nucleotide positions 1525-1527 (Figure 59). The predicted polypeptide precursor is 476 amino acids long (Figure 60). Clone DNA33206-1165 has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209372.

25 Analysis of the amino acid sequence of the full-length PRO223 polypeptide suggests that it possesses significant homology to various serine carboxypeptidase proteins, thereby indicating that PRO223 may be a novel serine carboxypeptidase.

30 EXAMPLE 28: Isolation of cDNA Clones Encoding Human PRO235

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is herein designated "DNA30927". Based on the DNA30927 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO235.

35 A pair of PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-TGGAATACCGCCTCGCAG-3' (SEQ ID NO:171)

reverse PCR primer 5'-CTTCTGCCCTTGGAGAAGATGGC-3' (SEQ ID NO:172)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA30927 sequence which had the following nucleotide sequence

hybridization probe

5'-GGACTCACTGGCCCAGGCCTCAATATCACCAAGCCAGGACGAT-3' (SEQ ID NO:173)

5 In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO235 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal liver tissue.

10 DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO235 [herein designated as DNA35558-1167] (SEQ ID NO:169) and the derived protein sequence for PRO235.

15 The entire nucleotide sequence of DNA35558-1167 is shown in Figure 61 (SEQ ID NO:169). Clone DNA35558-1167 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 667-669 and ending at the stop codon at nucleotide positions 2323-2325 (Figure 61). The predicted polypeptide precursor is 552 amino acids long (Figure 62). Clone DNA35558-1167 has been deposited with ATCC and is assigned ATCC deposit no. 209374.

20 Analysis of the amino acid sequence of the full-length PRO235 polypeptide suggests that portions of it possess significant homology to the human, mouse and *Xenopus* plexin protein, thereby indicating that PRO235 may be a novel plexin protein.

20

EXAMPLE 29: Isolation of cDNA Clones Encoding Human PRO236 and Human PRO262

Consensus DNA sequences were assembled relative to other EST sequences using phrap as described in Example 1 above. These consensus sequences are herein designated DNA30901 and DNA30847. Based on the DNA30901 and DNA30847 consensus sequences, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO236 and PRO262, respectively.

Based upon the DNA30901 consensus sequence, a pair of PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-TGGCTACTCCAAGACCCCTGGCATG-3' (SEQ ID NO:178)

30 reverse PCR primer 5'-TGGACAAATCCCCTGCTCAGCCC-3' (SEQ ID NO:179)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA30901 sequence which had the following nucleotide sequence

hybridization probe

5'-GGGCTTCACCGAACAGCAGTGGACCTTATTTGACCACCTGATGTCCAGGG-3' (SEQ ID NO:180)

35 Based upon the DNA30847 consensus sequence, a pair of PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-CCAGCTATGACTATGATGCACC-3' (SEQ ID NO:181)

reverse PCR primer 5'-TGGCACCCAGAATGGTGTGGCTC-3' (SEQ ID NO:182)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA30847 sequence which had the following nucleotide sequence

hybridization probe

5'-CGAGATGTCATCAGCAAGTTCCAGGAAGTTCTTGGACCTTACCTCC-3' (SEQ ID NO:183)

5 In order to screen several libraries for a source of full-length clones, DNA from the libraries was screened by PCR amplification with the PCR primer pairs identified above. Positive libraries were then used to isolate clones encoding the PRO236 and PRO262 genes using the probe oligonucleotides and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal lung tissue for PRO236 and 10 human fetal liver tissue for PRO262.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO236 [herein designated as DNA35599-1168] (SEQ ID NO:174), the derived protein sequence for PRO236, the full-length DNA sequence for PRO262 [herein designated as DNA36992-1168] (SEQ ID NO:176) and the derived protein sequence for PRO262.

15 The entire nucleotide sequence of DNA35599-1168 is shown in Figure 63 (SEQ ID NO:174). Clone DNA35599-1168 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 69-71 and ending at the stop codon at nucleotide positions 1977-1979 (Figure 63). The predicted polypeptide precursor is 636 amino acids long (Figure 64). Clone DNA35599-1168 has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209373.

20 The entire nucleotide sequence of DNA36992-1168 is shown in Figure 65 (SEQ ID NO:176). Clone DNA36992-1168 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 240-242 and ending at the stop codon at nucleotide positions 2202-2204 (Figure 65). The predicted polypeptide precursor is 654 amino acids long (Figure 66). Clone DNA36992-1168 has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209382.

25 Analysis of the amino acid sequence of the full-length PRO236 and PRO262 polypeptides suggests that portions of those polypeptides possess significant homology to β -galactosidase proteins derived from various sources, thereby indicating that PRO236 and PRO262 may be novel β -galactosidase homologs.

EXAMPLE 30: Isolation of cDNA Clones Encoding Human PRO239

30 A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is herein designated DNA30909. Based on the DNA30909 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO239.

35 A pair of PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-CCTCCCTCTATTACCCATGTC-3' (SEQ ID NO:186)

reverse PCR primer 5'-GACCAACTTCTCTGGGAGTGAGG-3' (SEQ ID NO:187)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA30909 sequence which had the following nucleotide sequence

hybridization probe

5'-GTCACCTTATTCTCTAACACAAAGCTCGAACCTTACCAAGTGGCAG-3'

(SEQ ID NO:188)

5 In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO239 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal lung tissue.

10 DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO239 [herein designated as DNA34407-1169] (SEQ ID NO:184) and the derived protein sequence for PRO239.

15 The entire nucleotide sequence of DNA34407-1169 is shown in Figure 67 (SEQ ID NO:184). Clone DNA34407-1169 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 72-74 and ending at the stop codon at nucleotide positions 1575-1577 (Figure 67). The predicted polypeptide precursor is 501 amino acids long (Figure 68). Clone DNA34407-1169 has been deposited with ATCC and is assigned ATCC deposit no.ATCC 209383.

20 Analysis of the amino acid sequence of the full-length PRO239 polypeptide suggests that portions of it possess significant homology to the densin protein, thereby indicating that PRO239 may be a novel molecule in the densin family.

20

EXAMPLE 31: Isolation of cDNA Clones Encoding Human PRO257

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is herein designated DNA28731. Based on the DNA28731 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO257.

25 A pair of PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-TCTCTATTCCAACTGTGGCG-3' (SEQ ID NO:191)

reverse PCR primer 5'-TTGATGACGATTCTGAAGGTGG-3' (SEQ ID NO:192)

30 Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA28731 sequence which had the following nucleotide sequence

hybridization probe

5'-GGAAGGATCCTCACCAAGCCCCAATTACCCAAAGCCGCATCCTGAGC-3' (SEQ ID NO:193)

35 In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO257 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal kidney tissue.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO257 [herein designated as DNA35841-1173 (SEQ ID NO:189) and the derived protein sequence for PRO257.

The entire nucleotide sequence of DNA35841-1173 is shown in Figure 69 (SEQ ID NO:189). Clone DNA35841-1173 contains a single open reading frame with an apparent translational initiation site at nucleotide 5 positions 964-966 and ending at the stop codon at nucleotide positions 2785-2787 (Figure 69). The predicted polypeptide precursor is 607 amino acids long (Figure 70). Clone DNA35841-1173 has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209403.

Analysis of the amino acid sequence of the full-length PRO257 polypeptide suggests that portions of it possess significant homology to the ebnerin protein, thereby indicating that PRO257 may be a novel protein 10 member related to the ebnerin protein.

EXAMPLE 32: Isolation of cDNA Clones Encoding Human PRO260

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is herein designated DNA30834. Based on the DNA30834 15 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO260.

PCR primers (forward and two reverse) were synthesized:

forward PCR primer: 5'-TGGTTTGACCAGGCCAAGTTCGG-3' (SEQ ID NO:196);

20 reverse PCR primer A: 5'-GGATTCTCATCCTCAAGGAAGAGCGG-3' (SEQ ID NO:197); and

reverse PCR primer B: 5'AACTTGCAGCATCAGCCACTCTGC-3' (SEQ ID NO:198)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA30834 sequence which had the following nucleotide sequence:

hybridization probe:

25 5'-TTCCGTGCCAGCTCGGTAGCGAGTGGTCTGGTGGTATTGGCA-3' (SEQ ID NO:199)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO260 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal kidney tissue.

30 DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO260 [herein designated as DNA33470-1175] (SEQ ID NO:194) and the derived protein sequence for PRO260.

The entire nucleotide sequence of DNA33470-1175 is shown in Figure 71 (SEQ ID NO:194). Clone DNA33470-1175 contains a single open reading frame with an apparent translational initiation site at nucleotide 35 positions 67-69 and ending at the stop codon 1468-1470 (see Figure 71). The predicted polypeptide precursor is 467 amino acids long (Figure 72). Clone DNA33470-1175 has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209398.

Analysis of the amino acid sequence of the full-length PRO260 polypeptide suggests that portions of it possess significant homology to the alpha-1-fucosidase precursor, thereby indicating that PRO260 may be a novel fucosidase.

EXAMPLE 33: Isolation of cDNA Clones Encoding Human PRO263

5 A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is herein designated DNA30914. Based on the DNA30914 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO263.

10 PCR primers (two forward and one reverse) were synthesized:

forward PCR primer 1: 5'-GAGCTTCCATCCAGGTGTCATGC-3' (SEQ ID NO:202);

forward PCR primer 2: 5'-GTCAGTGACAGTACCTACTCGG-3' (SEQ ID NO:203); reverse PCR primer:
5'-TGGAGCAGGAGGAGTAGTAGTAGG-3' (SEQ ID NO:204)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA30914

15 sequence which had the following nucleotide sequence:

hybridization probe:

5'-AGGAGGCCTGTAGGCTGCTGGACTAAGTTGGCCGGCAAGGACCAAGTT-3' (SEQ ID NO:205)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to 20 isolate clones encoding the PRO263 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal liver tissue.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO263 [herein designated as DNA34431-1177] (SEQ ID NO:200) and the derived protein sequence for PRO263.

25 The entire nucleotide sequence of DNA34431-1177 is shown in Figure 73 (SEQ ID NO:200). Clone DNA34431-1177 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 160-162 of SEQ ID NO:200 and ending at the stop codon after the nucleotide at position 1126-1128 of SEQ ID NO:200 (Figure 73). The predicted polypeptide precursor is 322 amino acids long (Figure 74). Clone DNA34431-1177 has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209399.

30 Analysis of the amino acid sequence of the full-length PRO263 polypeptide suggests that portions of it possess significant homology to CD44 antigen, thereby indicating that PRO263 may be a novel cell surface adhesion molecule.

EXAMPLE 34: Isolation of cDNA Clones Encoding Human PRO270

35 A consensus DNA sequence was assembled relative to the other identified EST sequences as described in Example 1 above, wherein the consensus sequence was designated herein as DNA35712. Based on the DNA35712 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that

contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO270. Forward and reverse PCR primers were synthesized:

forward PCR primer (.f1) 5'-GCTTGGATATCGCATGGGCCTAC-3' (SEQ ID NO:208)

forward PCR primer (.f2) 5'-TGGAGACAAATATCCCTGAGG-3' (SEQ ID NO:209)

reverse PCR primer (.r1) 5'-AACAGTTGCCACAGCATGGCAGG-3' (SEQ ID NO:210)

5 Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA35712 sequence which had the following nucleotide sequence

hybridization probe

5'-CCATTGATGAGGAACTAGAACGGACAAGAGGGTCACTGGATTGTGGAG-3'

(SEQ ID NO:211)

10 In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO270 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal lung tissue.

15 DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO270 [herein designated as DNA39510-1181] (SEQ ID NO:206) and the derived protein sequence for PRO270.

20 The entire nucleotide sequence of DNA39510-1181 is shown in Figure 75 (SEQ ID NO:206). Clone DNA39510-1181 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 3-5 and ending at the stop codon at nucleotide positions 891-893 (Fig. 75; SEQ ID NO:206). The predicted polypeptide precursor is 296 amino acids long (Fig. 76). Clone DNA39510-1181 has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209392.

Analysis of the amino acid sequence of the full-length PRO270 suggests that portions of it possess significant homology to the thioredoxin-protein, thereby indicating that the PRO270 protein may be a novel member of the thioredoxin family.

25

EXAMPLE 35: Isolation of cDNA Clones Encoding Human PRO271

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is herein designated DNA35737. Based on the DNA35737 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO271.

Forward and reverse PCR primers were synthesized:

forward PCR primer 1 5'-TGCTTCGCTACTGCCCTC-3' (SEQ ID NO:214)

forward PCR primer 2 5'-TTCCCTTGTGGGTTGGAG-3' (SEQ ID NO:215)

35 forward PCR primer 3 5'-AGGGCTGGAAGCCAGTTC-3' (SEQ ID NO:216)

reverse PCR primer 1 5'-AGCCAGTGAGGAAATGCG-3' (SEQ ID NO:217)

reverse PCR primer 2 5'-TGTCCAAAGTACACACACCTGAGG-3' (SEQ ID NO:218)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA35737 sequence which had the following nucleotide sequence
hybridization probe

5'-GATGCCACGATGCCAAGGTGGGACAGCTTGCCTGGAAAG-3' (SEQ ID NO:219)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was
5 screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to
isolate clones encoding the PRO271 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal brain tissue.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for
10 PRO271 [herein designated as DNA39423-1182] (SEQ ID NO:212) and the derived protein sequence for
PRO271.

The entire nucleotide sequence of DNA39423-1182 is shown in Figure 77 (SEQ ID NO:212). Clone
DNA39423-1182 contains a single open reading frame with an apparent translational initiation site at nucleotide
positions 101-103 and ending at the stop codon at nucleotide positions 1181-1183 (Figure 77). The predicted
polypeptide precursor is 360 amino acids long (Figure 78). Clone DNA39423-1182 has been deposited with
15 ATCC and is assigned ATCC deposit no. ATCC 209387.

Analysis of the amino acid sequence of the full-length PRO271 polypeptide suggests that it possess
significant homology to the proteoglycan link protein, thereby indicating that PRO271 may be a link protein
homolog.

20 EXAMPLE 36: Isolation of cDNA Clones Encoding Human PRO272

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described
in Example 1 above. This consensus sequence is herein designated DNA36460. Based on the DNA36460
consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained
the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for
25 PRO272.

Forward and reverse PCR primers were synthesized:

forward PCR primer (.f1) 5'-CGCAGGCCCTCATGGCCAGG-3' (SEQ ID NO:222)

forward PCR primer (.f2) 5'-GAAATCCTGGGTAATTGG-3' (SEQ ID NO:223)

reverse PCR primer 5'-GTGCGCGGTGCTCACAGCTCATC-3' (SEQ ID NO:224)

30 Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA36460
sequence which had the following nucleotide sequence

hybridization probe

5'-CCCCCTGAGCGACGCTCCCCATGATGACGCCACGGAACTTC-3' (SEQ ID NO:225)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was
35 screened by PCR amplification with the PCR primer pairs identified above. A positive library was then used
to isolate clones encoding the PRO272 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal lung tissue.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO272 [herein designated as DNA40620-1183] (SEQ ID NO:220) and the derived protein sequence for PRO272.

The entire nucleotide sequence of DNA40620-1183 is shown in Figure 79 (SEQ ID NO:220). Clone DNA40620-1183 contains a single open reading frame with an apparent translational initiation site at nucleotide 5 positions 35-37 and ending at the stop codon at nucleotide positions 1019-1021 (Figure 79). The predicted polypeptide precursor is 328 amino acids long (Figure 80). Clone DNA40620-1183 has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209388.

Analysis of the amino acid sequence of the full-length PRO272 polypeptide suggests that portions of it possess significant homology to the human and mouse reticulocalbin proteins, respectively, thereby indicating 10 that PRO272 may be a novel reticulocalbin protein.

EXAMPLE 37: Isolation of cDNA Clones Encoding Human PRO294

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is herein designated DNA35731. Based on the DNA35731 15 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO294.

Forward and reverse PCR primers were synthesized:

<u>forward PCR primer</u> (.f1)	5'-TGGTCTCGCACACCGATC-3'	(SEQ ID NO:228)
20 <u>forward PCR primer</u> (.f2)	5'-CTGCTGTCCACAGGGGAG-3'	(SEQ ID NO:229)
<u>forward PCR primer</u> (.f3)	5'-CCTTGAAGCATACTGCTC-3'	(SEQ ID NO:230)
<u>forward PCR primer</u> (.f4)	5'-GAGATAGCAATTCCGCC-3'	(SEQ ID NO:231)
<u>reverse PCR primer</u> (.r1)	5'-TTCCTCAAGAGGGCAGCC-3'	(SEQ ID NO:232)
<u>reverse PCR primer</u> (.r2)	5'-CTTGGCACCAATGTCCGAGATTTC-3'	
25 (SEQ ID NO:233)		

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA35731 sequence which had the following nucleotide sequence

hybridization probe

30 5'-GCTCTGAGGAAGGTGACCGCGGGGCCTCCGAACCCTGGCCTTG-3'
(SEQ ID NO:234)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pairs identified above. A positive library was then used to isolate clones encoding the PRO294 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal brain tissue.

35 DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO294 [herein designated as DNA40604-1187] (SEQ ID NO:226) and the derived protein sequence for PRO294.

The entire nucleotide sequence of DNA40604-1187 is shown in Figure 81 (SEQ ID NO:226). Clone DNA40604-1187 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 396-398 and ending at the stop codon at nucleotide positions 2046-2048 (Figure 81). The predicted polypeptide precursor is 550 amino acids long (Figure 82). Clone DNA40604-1187 has been deposited with ATCC and is assigned ATCC deposit no. 209394.

5 Analysis of the amino acid sequence of the full-length PRO294 polypeptide suggests that portions of it possess significant homology to portions of various collagen proteins, thereby indicating that PRO294 may be collagen-like molecule.

EXAMPLE 38: Isolation of cDNA Clones Encoding Human PRO295

10 A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is herein designated DNA35814. Based on the DNA35814 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO295.

15 Forward and reverse PCR primers were synthesized:

forward PCR primer (.f1) 5'-GCAGAGCGGAGATGCAGCGGTTG-3'

(SEQ ID NO:238)

forward PCR primer (.f2) 5'-CCCAGCATGTACTGCCAG-3' (SEQ ID NO:239)

forward PCR primer (.f3) 5'-TTGGCAGCTTCATGGAGG-3' (SEQ ID NO:240)

20 forward PCR primer (.f4) 5'-CCTGGCAAAATGCAAC-3' (SEQ ID NO:241)

reverse PCR primer (.r1) 5'-CTCCAGCTCCTGGCGCACCTCCTC-3' (SEQ ID NO:242)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA35814 sequence which had the following nucleotide sequence

hybridization probe

25 5'-GGCTCTCAGCTACCGCGCAGGAGCGAGGCCACCTCAATGAGATG-3'

(SEQ ID NO:243)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pairs identified above. A positive library was then used to isolate clones encoding the PRO295 gene using the probe oligonucleotide and one of the PCR primers.

30 RNA for construction of the cDNA libraries was isolated from human fetal lung tissue.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO295 [herein designated as DNA38268-1188] (SEQ ID NO:235) and the derived protein sequence for PRO295.

35 The entire nucleotide sequence of DNA38268-1188 is shown in Figure 83 (SEQ ID NO:235). Clone DNA38268-1188 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 153-155 and ending at the stop codon at nucleotide positions 1202-1204 (Figure 83). The predicted polypeptide precursor is 350 amino acids long (Figure 84). Clone DNA38268-1188 has been deposited with

ATCC and is assigned ATCC deposit no. 209421.

Analysis of the amino acid sequence of the full-length PRO295 polypeptide suggests that portions of it possess significant homology to the integrin proteins, thereby indicating that PRO295 may be a novel integrin.

EXAMPLE 39: Isolation of cDNA Clones Encoding Human PRO293

5 The extracellular domain (ECD) sequences (including the secretion signal, if any) of from about 950 known secreted proteins from the Swiss-Prot public protein database were used to search expressed sequence tag (EST) databases. The EST databases included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ™, Incyte Pharmaceuticals, Palo Alto, CA). The search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)) as a 10 comparison of the ECD protein sequences to a 6 frame translation of the EST sequence. Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into consensus DNA sequences with the program "phrap" (Phil Green, University of Washington, Seattle, Washington; <http://bozeman.mbt.washington.edu/phrap.docs/phrap.html>).

15 Based on an expression tag sequence designated herein as T08294 identified in the above analysis, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO293.

A pair of PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-AACAAAGGTAAGATGCCATCCTG-3' (SEQ ID NO:246)

reverse PCR primer 5'-AAACTTGTGATGGAGACCAGCTC-3' (SEQ ID NO:247)

20 Additionally, a synthetic oligonucleotide hybridization probe was constructed from the expression sequence tag which had the following nucleotide sequence

hybridization probe

5'-AGGGGCTGCAAAGCCTGGAGAGCCTCTCCTTCTATGACAACCAGC-3' (SEQ ID NO:248)

25 In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO293 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal brain tissue.

30 DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO293 [herein designated as DNA37151-1193] (SEQ ID NO:244) and the derived protein sequence for PRO293.

The entire nucleotide sequence of DNA37151-1193 is shown in Figure 85 (SEQ ID NO:244). Clone DNA37151-1193 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 881-883 and ending at the stop codon after nucleotide position 3019 of SEQ ID NO:244, Figure 85). The predicted polypeptide precursor is 713 amino acids long (Figure 86). Clone DNA37151-1193 has been 35 deposited with ATCC and is assigned ATCC deposit no. ATCC 209393.

Analysis of the amino acid sequence of the full-length PRO293 polypeptide suggests that portions of it possess significant homology to the NLRR proteins, thereby indicating that PRO293 may be a novel NLRR

protein.

EXAMPLE 40: Isolation of cDNA Clones Encoding Human PRO247

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is herein designated DNA33480. Based on the DNA33480 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO247.

A pair of PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-CAACAATGAGGGCACCAAGC-3' (SEQ ID NO:251)

10 reverse PCR primer 5'-GATGGCTAGGTTCTGGAGGTTCTG-3' (SEQ ID NO:252)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the DNA33480 expression sequence tag which had the following nucleotide sequence

hybridization probe

5'-CAACCTGCAGGAGATTGACCTCAAGGACAACAACCTCAAGACCATCG-3' (SEQ ID NO:253)

15 In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO247 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal brain tissue.

20 DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO247 [herein designated as DNA35673-1201] (SEQ ID NO:249) and the derived protein sequence for PRO247.

25 The entire nucleotide sequence of DNA35673-1201 is shown in Figure 89 (SEQ ID NO:249). Clone DNA35673-1201 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 80-82 of SEQ ID NO:249 and ending at the stop codon after nucleotide position 1717 of SEQ ID NO:249 (Figure 89). The predicted polypeptide precursor is 546 amino acids long (Figure 88). Clone DNA35673-1201 has been deposited with ATCC and is assigned ATCC deposit no. 209418.

30 Analysis of the amino acid sequence of the full-length PRO247 polypeptide suggests that portions of it possess significant homology to the densin molecule and KIAA0231, thereby indicating that PRO247 may be a novel leucine rich repeat protein.

EXAMPLE 41: Isolation of cDNA Clones Encoding Human PRO302, PRO303, PRO304, PRO307 and PRO343

35 Consensus DNA sequences were assembled relative to other EST sequences using phrap as described in Example 1 above. These consensus sequences are herein designated DNA35953, DNA35955, DNA35958, DNA37160 and DNA30895. Based on the DNA35953 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO302.

PCR primers (forward and reverse) were synthesized:

forward PCR primer 1 5'-GTCCGCAAGGATGCCTACATGTTC-3' (SEQ ID NO:264)

forward PCR primer 2 5'-GCAGAGGTGTCTAAGGTTG-3' (SEQ ID NO:265)

reverse PCR primer 5'-AGCTCTAGACCAATGCCAGCTTCC-3' (SEQ ID NO:266)

Also, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA35953 sequence

5 which had the following nucleotide sequence

hybridization probe

5'-GCCACCAACTCCTGCAAGAACCTCTCAGAACTGCCCTGGTCATG-3' (SEQ ID NO:267)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pairs identified above. A positive library was then used 10 to isolate clones encoding the PRO302 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal kidney tissue (LIB228).

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO302 [herein designated as DNA40370-1217] (SEQ ID NO:254) and the derived protein sequence for PRO302.

15 The entire nucleotide sequence of DNA40370-1217 is shown in Figure 89 (SEQ ID NO:254). Clone DNA40370-1217 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 34-36 and ending at the stop codon at nucleotide positions 1390-1392 (Figure 89). The predicted polypeptide precursor is 452 amino acids long (Figure 90). Various unique aspects of the PRO302 protein are shown in Figure 90. Clone DNA40370-1217 has been deposited with the ATCC on November 21, 1997 and 20 is assigned ATCC deposit no. ATCC 209485.

Based on the DNA35955 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO303.

A pair of PCR primers (forward and reverse) were synthesized:

25 forward PCR primer 5'-GGGGAATTCCACCCTATGACATTGCC-3' (SEQ ID NO:268)

reverse PCR primer 5'-GAATGCCCTGCAAGCATCAACTGG-3' (SEQ ID NO:269)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA35955 sequence which had the following nucleotide sequence:

hybridization probe

30 5'-GCACCTGTCACCTACACTAAACACATCCAGCCCATCTGTCTCCAGGCCTC-3' (SEQ ID NO:270)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pairs identified above. A positive library was then used to isolate clones encoding the PRO303 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal lung tissue (LIB25).

35 DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO303 [herein designated as DNA42551-1217] (SEQ ID NO:256) and the derived protein sequence for PRO303.

The entire nucleotide sequence of DNA42551-1217 is shown in Figure 91 (SEQ ID NO:256). Clone DNA42551-1217 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 20-22 and ending at the stop codon at nucleotide positions 962-964 (Figure 91). The predicted polypeptide precursor is 314 amino acids long (Figure 92). Various unique aspects of the PRO303 protein are shown in Figure 92. Clone DNA42551-1217 has been deposited on November 21, 1997 with the ATCC and 5 is assigned ATCC deposit no. ATCC 209483.

Based on the DNA35958 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO304.

Pairs of PCR primers (forward and reverse) were synthesized:

10 forward PCR primer 1 5'-GCGGAAGGGCAGAAATGGGACTCCAAG-3' (SEQ ID NO:271)
forward PCR primer 2 5'-CAGCCCTGCCACATGTGC-3' (SEQ ID NO:272)
forward PCR primer 3 5'-TACTGGGTGGTCAGCAAC-3' (SEQ ID NO:273)
reverse PCR primer 5'-GGCGAAGAGCAGGGTGAGACCCCCG-3' (SEQ ID NO:274)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA35958 15 sequence which had the following nucleotide sequence

hybridization probe

5'-GCCCTCATCCTCTGGCAAATGCAGTTACAGCCGGAGCCCGAC-3' (SEQ ID NO:275)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was 20 screened by PCR amplification with the PCR primer pairs identified above. A positive library was then used to isolate clones encoding the PRO304 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from 22 week human fetal brain tissue (LIB153).

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for 25 PRO304 [herein designated as DNA39520-1217] (SEQ ID NO:258) and the derived protein sequence for PRO304.

The entire nucleotide sequence of DNA39520-1217 is shown in Figure 93 (SEQ ID NO:258). Clone DNA39520-1217 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 34-36 and ending at the stop codon at nucleotide positions 1702-1704 (Figure 93). The predicted polypeptide precursor is 556 amino acids long (Figure 94). Various unique aspects of the PRO304 protein are 30 shown in Figure 94. Clone DNA39520-1217 has been deposited with ATCC on November 21, 1997 and is assigned ATCC deposit no. ATCC 209482.

Based on the DNA37160 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO307.

35 Pairs of PCR primers (forward and reverse) were synthesized:

forward PCR primer 1 5'-GGGCAGGGATTCCAGGGCTCC-3' (SEQ ID NO:276)
forward PCR primer 2 5'-GGCTATGACAGCAGGTT-3' (SEQ ID NO:277)

forward PCR primer 3 5'-TGACAATGACCGACCAGG-3' (SEQ ID NO:278)

reverse PCR primer 5'-GCATCGCATTGCTGGTAGAGCAAG-3' (SEQ ID NO:279)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA^{37/60} sequence which had the following nucleotide sequence

hybridization probe

5 5'-TTACAGTCCCCCTGGAAACCCACTTGGCCTGCATACCGCCTCCC-3' (SEQ ID NO:280)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pairs identified above. A positive library was then used to isolate clones encoding the PRO307 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal liver tissue (LIB229).

10 DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO307 [herein designated as DNA41225-1217] (SEQ ID NO:260) and the derived protein sequence for PRO307.

15 The entire nucleotide sequence of DNA41225-1217 is shown in Figure 95 (SEQ ID NO:260). Clone DNA41225-1217 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 92-94 and ending at the stop codon at nucleotide positions 1241-1243 (Figure 95). The predicted polypeptide precursor is 383 amino acids long (Figure 96). Various unique aspects of the PRO307 protein are shown in Figure 96. Clone DNA41225-1217 has been deposited with ATCC on November 21, 1997 and is assigned ATCC deposit no. ATCC 209491.

20 Based on the DNA30895 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO343.

A pair of PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-CGTCTCGAGCGCTCCATACAGTTCCCTTGCCCCA-3' (SEQ ID NO:281)

reverse PCR primer

25 5'-TGGAGGGGGAGCGGGATGCTTGTCTGGCGACTCCGGGGGCC
CCCTCATGTGCCAGGTGGA-3' (SEQ ID NO:282)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA30895 sequence which had the following nucleotide sequence

hybridization probe

30 5'-CCCTCAGACCCTGCAGAAGCTGAAGGTTCTATCATCGAC
TCGGAAGTCTGCAGCCATCTGTACTGGCGGGAGCAGGACAGGGACCCATCACTGAGGACATGC
TGTGTGCCGGCTACT-3' (SEQ ID NO:283)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pairs identified above. A positive library was then used to isolate clones encoding the PRO343 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal lung tissue (LIB26).

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for

PRO343 [herein designated as DNA43318-1217] (SEQ ID NO:262) and the derived protein sequence for PRO343.

The entire nucleotide sequence of DNA43318-1217 is shown in Figure 97 (SEQ ID NO:262). Clone DNA43318-1217 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 53-55 and ending at the stop codon at nucleotide positions 1004-1006 (Figure 97). The predicted 5 polypeptide precursor is 317 amino acids long (Figure 98). Various unique aspects of the PRO343 protein are shown in Figure 98. Clone DNA43318-1217 has been deposited with ATCC on November 21, 1997 and is assigned ATCC deposit no. ATCC 209481.

EXAMPLE 42: Isolation of cDNA Clones Encoding Human PRO328

10 A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is herein designated DNA35615. Based on the DNA35615 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO328.

15 Forward and reverse PCR primers were synthesized:

forward PCR primer 5'-TCCTGCAGTTCTGATGC-3' (SEQ ID NO:286)

reverse PCR primer 5'-CTCATATTGCACACCAGTAATCG-3' (SEQ ID NO:287)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA35615 sequence which had the following nucleotide sequence

20 hybridization probe

5'-ATGAGGAGAACGTTGATGGTGGAGCTGCACAACCTCTACCGGG-3'

(SEQ ID NO:288)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to 25 isolate clones encoding the PRO328 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal kidney tissue.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO328 [herein designated as DNA40587-1231] (SEQ ID NO:284) and the derived protein sequence for PRO328.

30 The entire nucleotide sequence of DNA40587-1231 is shown in Figure 99 (SEQ ID NO:284). Clone DNA40587-1231 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 15-17 and ending at the stop codon at nucleotide positions 1404-1406 (Figure 99). The predicted polypeptide precursor is 463 amino acids long (Figure 100). Clone DNA40587-1231 has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209438.

35 Analysis of the amino acid sequence of the full-length PRO328 polypeptide suggests that portions of it possess significant homology to the human glioblastoma protein and to the cysteine rich secretory protein thereby indicating that PRO328 may be a novel glioblastoma protein or cysteine rich secretory protein.

EXAMPLE 43: Isolation of cDNA Clones Encoding Human PRO335, PRO331 or PRO326

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is herein designated DNA36685. Based on the DNA36685 consensus sequence, and Incyte EST sequence no. 2228990, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the 5 full-length coding sequence for PRO335, PRO331 or PRO326.

Forward and reverse PCR primers were synthesized for the determination of PRO335:

<u>forward PCR primer</u>	5'-GGAACCGAATCTCAGCTA-3'	(SEQ ID NO:295)
<u>forward PCR primer</u>	5'-CCTAAACTGAACCTGGACCA-3'	(SEQ ID NO:296)
<u>forward PCR primer</u>	5'-GGCTGGAGACACTGAACCT-3'	(SEQ ID NO:297)
10 <u>forward PCR primer</u>	5'-ACAGCTGCACAGCTCAGAACAGTG-3'	(SEQ ID NO:298)
<u>reverse PCR primer</u>	5'-CATTCCCAGTATAAAAATTTTC-3'	(SEQ ID NO:299)
<u>reverse PCR primer</u>	5'-GGGTCTTGGTGAATGAGG-3'	(SEQ ID NO:300)
<u>reverse PCR primer</u>	5'-GTGCCTCTCGGTTACCAACCAATGG-3'	(SEQ ID NO:301)

Additionally, a synthetic oligonucleotide hybridization probe was constructed for the determination of PRO335

15 which had the following nucleotide sequence

hybridization probe

5'-GCGGCCACTGTTGGACCGAACTGTAACCAAGGGAGAACAGCCGTCTAC-3'
(SEQ ID NO:302)

Forward and reverse PCR primers were synthesized for the determination of PRO331:

20 <u>forward PCR primer</u>	5'-GCCTTGACAAACCTTCAGTCACTAGTGG-3'	(SEQ ID NO:303)
<u>reverse PCR primer</u>	5'-CCCCATGTGTCCATGACTGTTCCC-3'	(SEQ ID NO:304)

Additionally, a synthetic oligonucleotide hybridization probe was constructed for the determination of PRO331

which had the following nucleotide sequence

hybridization probe

25 5'-TACTGCCTCATGACCTCTTCACTCCCTGCATCATCTTAGAGCGG-3'
(SEQ ID NO:305)

Forward and reverse PCR primers were synthesized for the determination of PRO326:

<u>forward PCR primer</u>	5'-ACTCCAAGGAAATCGGATCCGTTTC-3'	(SEQ ID NO:306)
<u>reverse PCR primer</u>	5'-TTAGCAGCTGAGGATGGGCACAAC-3'	(SEQ ID NO:307)

30 Additionally, a synthetic oligonucleotide hybridization probe was constructed for the determination of PRO331
which had the following nucleotide sequence

hybridization probe

5'-GCCTTCACTGGTTGGATGCATTGGAGCATCTAGACCTGAGTGACAACGC-3'
(SEQ ID NO:308)

35 In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pairs identified above. A positive library was then used to isolate clones encoding the PRO335, PRO331 or PRO326 gene using the probe oligonucleotide and one of

the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal kidney tissue (PRO335 and PRO326) and human fetal brain (PRO331).

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO335, PRO331 or PRO326 [herein designated as SEQ ID NOS:289, 291 and 293, respectively; see Figures 5 103, 105 and 107, respectively], and the derived protein sequence for PRO335, PRO331 or PRO326 (see Figures 104, 106 and 108, respectively; SEQ ID NOS:290, 292 and 294, respectively).

The entire nucleotide sequences are shown in Figures 103, 105 and 107, deposited with the ATCC on June 2, 1998, November 7, 1997 and November 21, 1997, respectively.

Analysis of the amino acid sequence of the full-length PRO335, PRO331 or PRO326 polypeptide 10 suggests that portions of it possess significant homology to the LIG-1 protein, thereby indicating that PRO335, PRO331 and PRO326 may be a novel LIG-1-related protein.

EXAMPLE 44: Isolation of cDNA clones Encoding Human PRO332

Based upon an ECD homology search performed as described in Example 1 above, a consensus DNA 15 sequence designated herein as DNA36688 was assembled. Based on the DNA36688 consensus sequence, oligonucleotides were synthesized to identify by PCR a cDNA library that contained the sequence of interest and for use as probes to isolate a clone of the full-length coding sequence for PRO332.

A pair of PCR primers (forward and reverse) were synthesized:

5'-GCATTGGCCGCGAGACTTGCC-3' (SEQ ID NO:311)
20 5'-GCGGCCACGGTCTTGGAAATG-3' (SEQ ID NO:312)

A probe was also synthesized:

5'-TGGAGGAGCTAACCTCAGCTACAACCGCATACCAGCCCACAGG-3'
(SEQ ID NO:313)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was 25 screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO332 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from a human fetal liver library (LIB229).

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for DNA40982-1235 and the derived protein sequence for PRO332.

30 The entire nucleotide sequence of DNA40982-1235 is shown in Figure 109 (SEQ ID NO:309). Clone DNA40982-1235 contains a single open reading frame (with an apparent translational initiation site at nucleotide positions 342-344, as indicated in Figure 109). The predicted polypeptide precursor is 642 amino acids long, and has a calculated molecular weight of 72,067 (pI: 6.60). Clone DNA40982-1235 has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209433.

35 Based on a BLAST and FastA sequence alignment analysis of the full-length sequence, PRO332 shows about 30-40% amino acid sequence identity with a series of known proteoglycan sequences, including, for example, fibromodulin and fibromodulin precursor sequences of various species (FMOD_BOVIN, FMOD

CHICK, FMOD_RAT, FMOD_MOUSE, FMOD_HUMAN, P_R36773), osteomodulin sequences (AB000114_1, AB007848_1), decorin sequences (CFU83141_1, OCU03394_1, P_R42266, P_R42267, P_R42260, P_R89439), keratan sulfate proteoglycans (BTU48360_1, AF022890_1), corneal proteoglycan (AF022256_1), and bone/cartilage proteoglycans and proteoglycane precursors (PGS1_BOVIN, PGS2_MOUSE, PGS2_HUMAN).

5 **EXAMPLE 45: Isolation of cDNA clones Encoding Human PRO334**

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. Based on the consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO334.

10 Forward and reverse PCR primers were synthesized for the determination of PRO334:

forward PCR primer 5'-GATGGTCCTGCTCAAGTGCCCTG-3' (SEQ ID NO:316)

reverse PCR primer 5'-TTGCACTTGTAGGACCCACGTACG-3' (SEQ ID NO:317)

Additionally, a synthetic oligonucleotide hybridization probe was constructed for the determination of PRO334 which had the following nucleotide sequence

15 hybridization probe

5'-CTGATGGAGGACCTGTAGATGTTGATGAATGTGCTACAGGAAGAGCC-3'

(SEQ ID NO:318)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to 20 isolate clones encoding the PRO334 gene using the probe oligonucleotide and one of the PCR primers.

Human fetal kidney cDNA libraries used to isolate the cDNA clones were constructed by standard methods using commercially available reagents such as those from Invitrogen, San Diego, CA.

25 DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO334 [herein designated as DNA41379-1236] (SEQ ID NO:314) and the derived protein sequence for PRO334.

The entire nucleotide sequence of DNA41379-1236 (also referred to as UNQ295) is shown in Figure 109 (SEQ ID NO:314). Clone DNA41379-1236 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 203-205 and ending at the stop codon at nucleotide positions 1730-1732 (Figure 109). The predicted polypeptide precursor is 509 amino acids long (Figure 110). Clone 30 DNA41379-1236 has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209488.

Analysis of the amino acid sequence of the full-length PRO334 polypeptide suggests that portions of it possess significant homology to the fibulin and fibrillin proteins, thereby indicating that PRO334 may be a novel member of the EGF protein family.

35 **EXAMPLE 46: Isolation of cDNA Clones Encoding Human PRO346**

A consensus DNA sequence was identified using phrap as described in Example 1 above. Specifically, this consensus sequence is herein designated DNA38240. Based on the DNA38240 consensus sequence,

oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length PRO346 coding sequence.

RNA for construction of the cDNA libraries was isolated from human fetal liver. The cDNA libraries used to isolated the cDNA clones were constructed by standard methods using commercially available reagents (e.g., Invitrogen, San Diego, CA; Clontech, etc.). The cDNA was primed with oligo dT containing a NotI site, linked with blunt to SalI hemikinased adaptors, cleaved with NotI, sized appropriately by gel electrophoresis, and cloned in a defined orientation into a suitable cloning vector (such as pRK8 or pRKD; pRK5B is a precursor of pRK5D that does not contain the SfiI site; see, Holmes et al., *Science*, 253:1278-1280 (1991)) in the unique XhoI and NotI sites.

5 A cDNA clone was sequenced in entirety. The entire nucleotide sequence of DNA44167-1243 is shown in Figure 111 (SEQ ID NO:319). Clone DNA44167-1243 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 64-66 (Fig. 113; SEQ ID NO:319). The predicted polypeptide precursor is 450 amino acids long. Clone DNA44167-1243 has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209434 (designation DNA44167-1243).

10 Based on a BLAST, BLAST-2 and FastA sequence alignment analysis (using the ALIGN computer program) of the full-length sequence, PRO346 shows amino acid sequence identity to carcinoembryonic antigen (28%).

15 The oligonucleotide sequences used in the above procedure were the following:

OLI2691 (38240.f1)

5'-GATCCTGTCACAAAGCCAGTGGTGC-3' (SEQ ID NO:321)

20 OLI2693 (38240.r1)

5'-CACTGACAGGGTTCCTCACCCAGG-3' (SEQ ID NO:322)

OLI2692 (38240.p1)

5'-CTCCCTCTGGGCTGTGGAGTATGTGGGGAACATGACCTGACATG-3' (SEQ ID NO:323)

25 EXAMPLE 47: Isolation of cDNA Clones Encoding Human PRO268

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is herein designated DNA35698. Based on the DNA35698 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for 30 PRO268.

Forward and reverse PCR primers were synthesized:

forward PCR primer 1 5'-TGAGGTGGCAAGCGCGAAATG-3' (SEQ ID NO:326)

forward PCR primer 2 5'-TATGTGGATCAGGACGTGCC-3' (SEQ ID NO:327)

forward PCR primer 3 5'-TGCAGGGTTCAGTCTAGATTG-3' (SEQ ID NO:328)

35 reverse PCR primer 5'-TTGAAGGACAAAGGCAATCTGCCAC-3' (SEQ ID NO:329)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA35698 sequence which had the following nucleotide sequence

hybridization probe

5'-GGAGTCTTGCAGTCCCTGGCAGTCCTGGTGCCTGCTTGGG-3' (SEQ ID NO:330)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO268 gene using the probe oligonucleotide and one of the PCR primers.

5 RNA for construction of the cDNA libraries was isolated from human fetal lung tissue.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO268 [herein designated as DNA39427-1179] (SEQ ID NO:324) and the derived protein sequence for PRO268.

The entire nucleotide sequence of DNA39427-1179 is shown in Figure 113 (SEQ ID NO:324). Clone 10 DNA39427-1179 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 13-15 and ending at the stop codon at nucleotide positions 853-855 (Figure 113). The predicted polypeptide precursor is 280 amino acids long (Figure 114). Clone DNA39427-1179 has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209395.

Analysis of the amino acid sequence of the full-length PRO268 polypeptide suggests that it possess 15 significant homology to protein disulfide isomerase, thereby indicating that PRO268 may be a novel protein disulfide isomerase.

EXAMPLE 48: Isolation of cDNA Clones Encoding Human PRO330

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described 20 in Example 1 above. This consensus sequence is herein designated DNA35730. Based on the DNA35730 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO330.

Forward and reverse PCR primers were synthesized:

25 forward PCR primer 1 5'-CCAGGCACAATTCCAGA-3' (SEQ ID NO:333)

forward PCR primer 2 5'-GGACCCCTCTGTGTGCCAG-3' (SEQ ID NO:334)

reverse PCR primer 1 5'-GGTCTCAAGAACTCCTGTC-3' (SEQ ID NO:335)

reverse PCR primer 2 5'-ACACTCAGCATTGCCTGGTACTTG-3' (SEQ ID NO:336)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus sequence which 30 had the following nucleotide sequence

hybridization probe

5'-GGGCACATGACTGACCTGATTATGCAGAGAAAGAGCTGGTGCAG-3' (SEQ ID NO:337)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to 35 isolate clones encoding the PRO330 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal liver tissue.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for

PRO330 [herein designated as DNA40603-1232] (SEQ ID NO:331) and the derived protein sequence for PRO330.

The entire nucleotide sequence of DNA40603-1232 is shown in Figure 115 (SEQ ID NO:331). Clone DNA40603-1232 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 167-169 and ending at the stop codon at nucleotide positions 1766-1768 (Figure 115). The predicted 5 polypeptide precursor is 533 amino acids long (Figure 116). Clone DNA40603-1232 has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209486 on November 21, 1997.

Analysis of the amino acid sequence of the full-length PRO330 polypeptide suggests that portions of it possess significant homology to the mouse prolyl 4-hydroxylase alpha subunit protein, thereby indicating that PRO330 may be a novel prolyl 4-hydroxylase alpha subunit polypeptide.

10

EXAMPLE 49: Isolation of cDNA Clones Encoding Human PRO310

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is herein designated DNA40553. Based on the DNA40553 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained 15 the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO310.

Forward and reverse PCR primers were synthesized:

forward PCR primer 1 5'-TCCCCAAGCCGTTCTAGACGCGG-3' (SEQ ID NO:342)

forward PCR primer 2 5'-CTGGTTCTTCCTTGCACG-3' (SEQ ID NO:343)

20 reverse PCR primer 5'-GCCAAATGCCCTAAGGCGGTATACCCC-3' (SEQ ID NO:344)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus sequence which had the following nucleotide sequence

hybridization probe

5'-GGGTGTGATGCTTGAAGCATTCTGTGCTTGATCACTATGCTAGGAC-3' (SEQ ID NO:345)

25 In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO310 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal liver tissue.

30 DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO310 [herein designated as DNA43046-1225 (SEQ ID NO:340) and the derived protein sequence for PRO310 (SEQ ID NO:341)].

The entire nucleotide sequence of DNA43046-1225 is shown in Figure 119 (SEQ ID NO:340). Clone DNA43046-1225 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 81-83 and ending at the stop codon at nucleotide positions 1035-1037 (Figure 119). The predicted 35 polypeptide precursor is 318 amino acids long (Figure 120) and has a calculated molecular weight of approximately 36,382 daltons. Clone DNA43046-1225 has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209484.

Analysis of the amino acid sequence of the full-length PRO310 polypeptide suggests that portions of it possess homology to *C. elegans* proteins and to fringe, thereby indicating that PRO310 may be involved in development.

EXAMPLE 50: Isolation of cDNA clones Encoding Human PRO339

5 An expressed sequence tag (EST) DNA database (LIFESEQ™, Incyte Pharmaceuticals, Palo Alto, CA) was searched and ESTs were identified. An assembly of Incyte clones and a consensus sequence was formed using phrap as described in Example 1 above.

Forward and reverse PCR primers were synthesized based upon the assembly-created consensus sequence:

10 forward PCR primer 1 5'-GGGATGCAGGTGGTGTCTCATGGGG-3' (SEQ ID NO:346)

forward PCR primer 2 5'-CCCTCATGTACCGGCTCC-3' (SEQ ID NO:347)

forward PCR primer 3 5'-GTGTGACACAGCGTGGGC-3' (SEQ ID NO:43)

forward PCR primer 4 5'-GACCGGCAGGCTTCTGCG-3' (SEQ ID NO:44)

reverse PCR primer 1 5'-CAGCAGCTTCAGCCACCAGGAGTGG-3' (SEQ ID NO:45)

15 reverse PCR primer 2 5'-CTGAGCCGTGGGCTGCAGTCTCGC-3' (SEQ ID NO:46)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus sequence which had the following nucleotide sequence

hybridization probe

5'-CCGACTACGACTGGTTCTTCATCATGCAGGATGACACATATGTGC-3' (SEQ ID NO:47)

20 In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pairs identified above. A positive library was then used to isolate clones encoding the PRO339 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal liver tissue.

25 A cDNA clone was sequenced in entirety. The entire nucleotide sequence of DNA43466-1225 is shown in Figure 117 (SEQ ID NO:338). Clone DNA43466-1225 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 333-335 and ending at the stop codon found at nucleotide positions 2649-2651 (Figure 117; SEQ ID NO:338). The predicted polypeptide precursor is 772 amino acids long and has a calculated molecular weight of approximately 86,226 daltons. Clone DNA43466-1225 has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209490.

30 Based on a BLAST and FastA sequence alignment analysis (using the ALIGN computer program) of the full-length sequence, PRO339 has homology to *C. elegans* proteins and collagen-like polymer sequences as well as to fringe, thereby indicating that PRO339 may be involved in development or tissue growth.

EXAMPLE 51: Isolation of cDNA Clones Encoding Human PRO244

35 A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. Based on this consensus sequence, oligonucleotides were synthesized to identify by PCR a cDNA library that contained the sequence of interest and for use as probes to isolate a clone of the full-length

coding sequence for PRO244.

A pair of PCR primers (forward and reverse) were synthesized:

5'-TTCAGCTTCTGGGATGTAGGG-3' (30923.f1) (SEQ ID NO:378)

5'-TATTCTACCATTCACAAATCCG-3' (30923.r1) (SEQ ID NO:379)

A probe was also synthesized:

5 5'-GGAGGACTGTGCCACCATGAGAGACTCTCAACCCAAGGCAAAATTGG-3' (30923.p1) (SEQ IDNO:380)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO244 gene using the probe oligonucleotide and one of the PCR primers.

10 RNA for construction of the cDNA libraries was isolated from a human fetal kidney library. DNA sequencing of the clones isolated as described above gave the full-length DNA sequence and the derived protein sequence for PRO244.

15 The entire nucleotide sequence of PRO244 is shown in Figure 121 (SEQ ID NO:376). Clone DNA35668-1171 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 106-108 (Fig. 121). The predicted polypeptide precursor is 219 amino acids long. Clone DNA35668-1171 has been deposited with ATCC (designated as DNA35663-1171) and is assigned ATCC deposit no. ATCC209371. The protein has a cytoplasmic domain (aa 1-20), a transmembrane domain (aa 21-46), and an extracellular domain (aa 47-219), with a C-lectin domain at aa 55-206.

20 Based on a BLAST and FastA sequence alignment analysis of the full-length sequence, PRO244 shows notable amino acid sequence identity to hepatic lectin gallus gallus (43%), HIC hp120-binding C-type lectin (42%), macrophage lectin 2 (HUMHML2-1, 41%), and sequence PR32188 (44%).

EXAMPLE 52: Use of PRO Polypeptide-Encoding Nucleic Acid as Hybridization Probes

25 The following method describes use of a nucleotide sequence encoding a PRO polypeptide as a hybridization probe.

DNA comprising the coding sequence of of a PRO polypeptide of interest as disclosed herein may be employed as a probe or used as a basis from which to prepare probes to screen for homologous DNAs (such as those encoding naturally-occurring variants of the PRO polypeptide) in human tissue cDNA libraries or human tissue genomic libraries.

30 Hybridization and washing of filters containing either library DNAs is performed under the following high stringency conditions. Hybridization of radiolabeled PRO polypeptide-encoding nucleic acid-derived probe to the filters is performed in a solution of 50% formamide, 5x SSC, 0.1% SDS, 0.1% sodium pyrophosphate, 50 mM sodium phosphate, pH 6.8, 2x Denhardt's solution, and 10% dextran sulfate at 42°C for 20 hours. Washing of the filters is performed in an aqueous solution of 0.1x SSC and 0.1% SDS at 42°C.

35 DNAs having a desired sequence identity with the DNA encoding full-length native sequence PRO polypeptide can then be identified using standard techniques known in the art.

EXAMPLE 53: Expression of PRO Polypeptides in *E. coli*

This example illustrates preparation of an unglycosylated form of a desired PRO polypeptide by recombinant expression in *E. coli*.

The DNA sequence encoding the desired PRO polypeptide is initially amplified using selected PCR primers. The primers should contain restriction enzyme sites which correspond to the restriction enzyme sites 5 on the selected expression vector. A variety of expression vectors may be employed. An example of a suitable vector is pBR322 (derived from *E. coli*; see Bolivar et al., *Gene*, 2:95 (1977)) which contains genes for ampicillin and tetracycline resistance. The vector is digested with restriction enzyme and dephosphorylated. The PCR amplified sequences are then ligated into the vector. The vector will preferably include sequences 10 which encode for an antibiotic resistance gene, a trp promoter, a polyhis leader (including the first six STII codons, polyhis sequence, and enterokinase cleavage site), the specific PRO polypeptide coding region, lambda transcriptional terminator, and an argU gene.

The ligation mixture is then used to transform a selected *E. coli* strain using the methods described in 15 Sambrook et al., *supra*. Transformants are identified by their ability to grow on LB plates and antibiotic resistant colonies are then selected. Plasmid DNA can be isolated and confirmed by restriction analysis and DNA sequencing.

Selected clones can be grown overnight in liquid culture medium such as LB broth supplemented with antibiotics. The overnight culture may subsequently be used to inoculate a larger scale culture. The cells are then grown to a desired optical density, during which the expression promoter is turned on.

20 After culturing the cells for several more hours, the cells can be harvested by centrifugation. The cell pellet obtained by the centrifugation can be solubilized using various agents known in the art, and the solubilized PRO polypeptide can then be purified using a metal chelating column under conditions that allow tight binding of the protein.

25 PRO187, PRO317, PRO301, PRO224 and PRO238 were successfully expressed in *E. coli* in a poly-His tagged form, using the following procedure. The DNA encoding PRO187, PRO317, PRO301, PRO224 or PRO238 was initially amplified using selected PCR primers. The primers contained restriction enzyme sites which correspond to the restriction enzyme sites on the selected expression vector, and other useful sequences providing for efficient and reliable translation initiation, rapid purification on a metal chelation column, and proteolytic removal with enterokinase. The PCR-amplified, poly-His tagged sequences were then ligated into 30 an expression vector, which was used to transform an *E. coli* host based on strain 52 (W3110 fuhA(tonA) lon galE rpoHts(htpRts) clpP(lacIq)). Transformants were first grown in LB containing 50 mg/ml carbenicillin at 30°C with shaking until an O.D.600 of 3-5 was reached. Cultures were then diluted 50-100 fold into CRAP media (prepared by mixing 3.57 g (NH₄)₂SO₄, 0.71 g sodium citrate·2H₂O, 1.07 g KCl, 5.36 g Difco yeast extract, 5.36 g Sheffield hycase SF in 500 mL water, as well as 110 mM MPOS, pH 7.3, 0.55% (w/v) glucose and 7 mM MgSO₄) and grown for approximately 20-30 hours at 30°C with shaking. Samples were removed 35 to verify expression by SDS-PAGE analysis, and the bulk culture is centrifuged to pellet the cells. Cell pellets were frozen until purification and refolding.

E. coli paste from 0.5 to 1 L fermentations (6-10 g pellets) was resuspended in 10 volumes (w/v) in 7 M guanidine, 20 mM Tris, pH 8 buffer. Solid sodium sulfite and sodium tetrathionate is added to make final concentrations of 0.1M and 0.02 M, respectively, and the solution was stirred overnight at 4°C. This step results in a denatured protein with all cysteine residues blocked by sulfitolization. The solution was centrifuged at 40,000 rpm in a Beckman Ultracentrifuge for 30 min. The supernatant was diluted with 3-5 volumes of metal 5 chelate column buffer (6 M guanidine, 20 mM Tris, pH 7.4) and filtered through 0.22 micron filters to clarify. Depending the clarified extract was loaded onto a 5 ml Qiagen Ni-NTA metal chelate column equilibrated in the metal chelate column buffer. The column was washed with additional buffer containing 50 mM imidazole (Calbiochem, Utrol grade), pH 7.4. The protein was eluted with buffer containing 250 mM imidazole. Fractions containing the desired protein were pooled and stored at 4°C. Protein concentration was estimated 10 by its absorbance at 280 nm using the calculated extinction coefficient based on its amino acid sequence.

The proteins were refolded by diluting sample slowly into freshly prepared refolding buffer consisting of: 20 mM Tris, pH 8.6, 0.3 M NaCl, 2.5 M urea, 5 mM cysteine, 20 mM glycine and 1 mM EDTA. Refolding volumes were chosen so that the final protein concentration was between 50 to 100 micrograms/ml. The refolding solution was stirred gently at 4°C for 12-36 hours. The refolding reaction was quenched by the 15 addition of TFA to a final concentration of 0.4% (pH of approximately 3). Before further purification of the protein, the solution was filtered through a 0.22 micron filter and acetonitrile was added to 2-10% final concentration. The refolded protein was chromatographed on a Poros R1/H reversed phase column using a mobile buffer of 0.1% TFA with elution with a gradient of acetonitrile from 10 to 80%. Aliquots of fractions with A280 absorbance were analyzed on SDS polyacrylamide gels and fractions containing homogeneous 20 refolded protein were pooled. Generally, the properly refolded species of most proteins are eluted at the lowest concentrations of acetonitrile since those species are the most compact with their hydrophobic interiors shielded from interaction with the reversed phase resin. Aggregated species are usually eluted at higher acetonitrile concentrations. In addition to resolving misfolded forms of proteins from the desired form, the reversed phase step also removes endotoxin from the samples.

25 Fractions containing the desired folded PRO187, PRO317, PRO301, PRO224 and PRO238 proteins, respectively, were pooled and the acetonitrile removed using a gentle stream of nitrogen directed at the solution. Proteins were formulated into 20 mM Hepes, pH 6.8 with 0.14 M sodium chloride and 4% mannitol by dialysis or by gel filtration using G25 Superfine (Pharmacia) resins equilibrated in the formulation buffer and sterile filtered.

30

EXAMPLE 54: Expression of PRO Polypeptides in Mammalian Cells

This example illustrates preparation of a glycosylated form of a desired PRO polypeptide by recombinant expression in mammalian cells.

The vector, pRK5 (see EP 307,247, published March 15, 1989), is employed as the expression vector.

35 Optionally, the PRO polypeptide-encoding DNA is ligated into pRK5 with selected restriction enzymes to allow insertion of the PRO polypeptide DNA using ligation methods such as described in Sambrook et al., supra. The resulting vector is called pRK5-PRO polypeptide.

In one embodiment, the selected host cells may be 293 cells. Human 293 cells (ATCC CCL 1573) are grown to confluence in tissue culture plates in medium such as DMEM supplemented with fetal calf serum and optionally, nutrient components and/or antibiotics. About 10 μ g pRK5-PRO polypeptide DNA is mixed with about 1 μ g DNA encoding the VA RNA gene [Thimmappaya et al., Cell, 31:543 (1982)] and dissolved in 500 μ l of 1 mM Tris-HCl, 0.1 mM EDTA, 0.227 M CaCl₂. To this mixture is added, dropwise, 500 μ l of 50 mM 5 HEPES (pH 7.35), 280 mM NaCl, 1.5 mM NaPO₄, and a precipitate is allowed to form for 10 minutes at 25°C. The precipitate is suspended and added to the 293 cells and allowed to settle for about four hours at 37°C. The culture medium is aspirated off and 2 ml of 20% glycerol in PBS is added for 30 seconds. The 293 cells are then washed with serum free medium, fresh medium is added and the cells are incubated for about 5 days.

Approximately 24 hours after the transfections, the culture medium is removed and replaced with culture 10 medium (alone) or culture medium containing 200 μ Ci/ml ³⁵S-cysteine and 200 μ Ci/ml ³⁵S-methionine. After a 12 hour incubation, the conditioned medium is collected, concentrated on a spin filter, and loaded onto a 15% SDS gel. The processed gel may be dried and exposed to film for a selected period of time to reveal the presence of PRO polypeptide. The cultures containing transfected cells may undergo further incubation (in serum free medium) and the medium is tested in selected bioassays.

15 In an alternative technique, PRO polypeptide may be introduced into 293 cells transiently using the dextran sulfate method described by Somparyrac et al., Proc. Natl. Acad. Sci., 78:7575 (1981). 293 cells are grown to maximal density in a spinner flask and 700 μ g pRK5-PRO polypeptide DNA is added. The cells are first concentrated from the spinner flask by centrifugation and washed with PBS. The DNA-dextran precipitate is incubated on the cell pellet for four hours. The cells are treated with 20% glycerol for 90 seconds, washed 20 with tissue culture medium, and re-introduced into the spinner flask containing tissue culture medium, 5 μ g/ml bovine insulin and 0.1 μ g/ml bovine transferrin. After about four days, the conditioned media is centrifuged and filtered to remove cells and debris. The sample containing expressed PRO polypeptide can then be concentrated and purified by any selected method, such as dialysis and/or column chromatography.

In another embodiment, PRO polypeptides can be expressed in CHO cells. The pRK5-PRO polypeptide 25 can be transfected into CHO cells using known reagents such as CaPO₄ or DEAE-dextran. As described above, the cell cultures can be incubated, and the medium replaced with culture medium (alone) or medium containing a radiolabel such as ³⁵S-methionine. After determining the presence of PRO polypeptide, the culture medium may be replaced with serum free medium. Preferably, the cultures are incubated for about 6 days, and then the conditioned medium is harvested. The medium containing the expressed PRO polypeptide can then be 30 concentrated and purified by any selected method.

Epitope-tagged PRO polypeptide may also be expressed in host CHO cells. The PRO polypeptide may be subcloned out of the pRK5 vector. The subclone insert can undergo PCR to fuse in frame with a selected epitope tag such as a poly-his tag into a Baculovirus expression vector. The poly-his tagged PRO polypeptide insert can then be subcloned into a SV40 driven vector containing a selection marker such as DHFR for selection 35 of stable clones. Finally, the CHO cells can be transfected (as described above) with the SV40 driven vector. Labeling may be performed, as described above, to verify expression. The culture medium containing the expressed poly-His tagged PRO polypeptide can then be concentrated and purified by any selected method, such

as by Ni^{2+} -chelate affinity chromatography.

PRO211, PRO217, PRO230, PRO219, PRO245, PRO221, PRO258, PRO301, PRO224, PRO222, PRO234, PRO229, PRO223, PRO328 and PRO332 were successfully expressed in CHO cells by both a transient and a stable expression procedure. In addition, PRO232, PRO265, PRO246, PRO228, PRO227, PRO220, PRO266, PRO269, PRO287, PRO214, PRO231, PRO233, PRO238, PRO244, PRO235, PRO236, PRO262, 5 PRO239, PRO257, PRO260, PRO263, PRO270, PRO271, PRO272, PRO294, PRO295, PRO293, PRO247, PRO303 and PRO268 were successfully transiently expressed in CHO cells.

Stable expression in CHO cells was performed using the following procedure. The proteins were expressed as an IgG construct (immunoadhesin), in which the coding sequences for the soluble forms (e.g. extracellular domains) of the respective proteins were fused to an IgG1 constant region sequence containing the 10 hinge, CH2 and CH2 domains and/or a poly-His tagged form.

Following PCR amplification, the respective DNAs were subcloned in a CHO expression vector using standard techniques as described in Ausubel et al., *Current Protocols of Molecular Biology*, Unit 3.16; John Wiley and Sons (1997). CHO expression vectors are constructed to have compatible restriction sites 5' and 3' of the DNA of interest to allow the convenient shuttling of cDNA's. The vector used expression in CHO cells 15 is as described in Lucas *et al.*, *Nucl. Acids Res.* **24**: 9 (1774-1779 (1996), and uses the SV40 early promoter/enhancer to drive expression of the cDNA of interest and dihydrofolate reductase (DHFR). DHFR expression permits selection for stable maintenance of the plasmid following transfection.

Twelve micrograms of the desired plasmid DNA were introduced into approximately 10 million CHO cells using commercially available transfection reagents Superfect[®] (Qiagen), Dospel[®] or Fugene[®] (Boehringer 20 Mannheim). The cells were grown and described in Lucas *et al.*, *supra*. Approximately 3×10^7 cells are frozen in an ampule for further growth and production as described below.

The ampules containing the plasmid DNA were thawed by placement into water bath and mixed by vortexing. The contents were pipetted into a centrifuge tube containing 10 mLs of media and centrifuged at 1000 rpm for 5 minutes. The supernatant was aspirated and the cells were resuspended in 10 mL of selective media 25 (0.2 μm filtered PS20 with 5% 0.2 μm diafiltered fetal bovine serum). The cells were then aliquoted into a 100 mL spinner containing 90 mL of selective media. After 1-2 days, the cells were transferred into a 250 mL spinner filled with 150 mL selective growth medium and incubated at 37°C. After another 2-3 days, a 250 mL, 500 mL and 2000 mL spinners were seeded with 3×10^5 cells/mL. The cell media was exchanged with fresh media by centrifugation and resuspension in production medium. Although any suitable CHO media may be 30 employed, a production medium described in US Patent No. 5,122,469, issued June 16, 1992 was actually used. 3L production spinner is seeded at 1.2×10^6 cells/mL. On day 0, the cell number pH were determined. On day 1, the spinner was sampled and sparging with filtered air was commenced. On day 2, the spinner was sampled, the temperature shifted to 33°C, and 30 mL of 500 g/L glucose and 0.6 mL of 10% antifoam (e.g., 35% polydimethylsiloxane emulsion, Dow Corning 365 Medical Grade Emulsion). Throughout the production, pH 35 was adjusted as necessary to keep at around 7.2. After 10 days, or until viability dropped below 70%, the cell culture was harvested by centrifugation and filtering through a 0.22 μm filter. The filtrate was either stored at 4°C or immediately loaded onto columns for purification.

For the poly-His tagged constructs, the proteins were purified using a Ni-NTA column (Qiagen). Before purification, imidazole was added to the conditioned media to a concentration of 5 mM. The conditioned media was pumped onto a 6 ml Ni-NTA column equilibrated in 20 mM Hepes, pH 7.4, buffer containing 0.3 M NaCl and 5 mM imidazole at a flow rate of 4-5 ml/min. at 4°C. After loading, the column was washed with additional equilibration buffer and the protein eluted with equilibration buffer containing 0.25 M imidazole. The 5 highly purified protein was subsequently desalted into a storage buffer containing 10 mM Hepes, 0.14 M NaCl and 4% mannitol, pH 6.8, with a 25 ml G25 Superfine (Pharmacia) column and stored at -80°C.

Immunoadhesin (Fc containing) constructs were purified from the conditioned media as follows. The conditioned medium was pumped onto a 5 ml Protein A column (Pharmacia) which had been equilibrated in 20 mM Na phosphate buffer, pH 6.8. After loading, the column was washed extensively with equilibration buffer 10 before elution with 100 mM citric acid, pH 3.5. The eluted protein was immediately neutralized by collecting 1 ml fractions into tubes containing 275 µL of 1 M Tris buffer, pH 9. The highly purified protein was subsequently desalted into storage buffer as described above for the poly-His tagged proteins. The homogeneity was assessed by SDS polyacrylamide gels and by N-terminal amino acid sequencing by Edman degradation.

PRO211, PRO217, PRO230, PRO232, PRO187, PRO265, PRO219, PRO246, PRO228, PRO533, 15 PRO245, PRO221, PRO227, PRO220, PRO258, PRO266, PRO269, PRO287, PRO214, PRO317, PRO301, PRO224, PRO222, PRO234, PRO231, PRO229, PRO233, PRO238, PRO223, PRO235, PRO236, PRO262, PRO239, PRO257, PRO260, PRO263, PRO270, PRO271, PRO272, PRO294, PRO295, PRO293, PRO247, PRO304, PRO302, PRO307, PRO303, PRO343, PRO328, PRO326, PRO331, PRO332, PRO334, PRO346, PRO268, PRO330, PRO310 and PRO339 were also successfully transiently expressed in COS cells.

20

EXAMPLE 55: Expression of PRO Polypeptides in Yeast

The following method describes recombinant expression of a desired PRO polypeptide in yeast.

First, yeast expression vectors are constructed for intracellular production or secretion of PRO polypeptides from the ADH2/GAPDH promoter. DNA encoding a desired PRO polypeptide, a selected signal 25 peptide and the promoter is inserted into suitable restriction enzyme sites in the selected plasmid to direct intracellular expression of the PRO polypeptide. For secretion, DNA encoding the PRO polypeptide can be cloned into the selected plasmid, together with DNA encoding the ADH2/GAPDH promoter, the yeast alpha-factor secretory signal/leader sequence, and linker sequences (if needed) for expression of the PRO polypeptide.

Yeast cells, such as yeast strain AB110, can then be transformed with the expression plasmids described 30 above and cultured in selected fermentation media. The transformed yeast supernatants can be analyzed by precipitation with 10% trichloroacetic acid and separation by SDS-PAGE, followed by staining of the gels with Coomassie Blue stain.

Recombinant PRO polypeptide can subsequently be isolated and purified by removing the yeast cells from the fermentation medium by centrifugation and then concentrating the medium using selected cartridge 35 filters. The concentrate containing the PRO polypeptide may further be purified using selected column chromatography resins.

EXAMPLE 56: Expression of PRO Polypeptides in Baculovirus-Infected Insect Cells

The following method describes recombinant expression of PRO polypeptides in Baculovirus-infected insect cells.

The desired PRO polypeptide is fused upstream of an epitope tag contained with a baculovirus expression vector. Such epitope tags include poly-his tags and immunoglobulin tags (like Fc regions of IgG).

5 A variety of plasmids may be employed, including plasmids derived from commercially available plasmids such as pVL1393 (Novagen). Briefly, the PRO polypeptide or the desired portion of the PRO polypeptide (such as the sequence encoding the extracellular domain of a transmembrane protein) is amplified by PCR with primers complementary to the 5' and 3' regions. The 5' primer may incorporate flanking (selected) restriction enzyme sites. The product is then digested with those selected restriction enzymes and subcloned into the expression

10 vector.

Recombinant baculovirus is generated by co-transfected the above plasmid and BaculoGold™ virus DNA (Pharmingen) into *Spodoptera frugiperda* ("Sf9") cells (ATCC CRL 1711) using lipofectin (commercially available from GIBCO-BRL). After 4-5 days of incubation at 28°C, the released viruses are harvested and used for further amplifications. Viral infection and protein expression is performed as described by O'Reilly et al.,

15 Baculovirus expression vectors: A laboratory Manual, Oxford: Oxford University Press (1994).

Expressed poly-his tagged PRO polypeptide can then be purified, for example, by Ni²⁺-chelate affinity chromatography as follows. Extracts are prepared from recombinant virus-infected Sf9 cells as described by Rupert et al., *Nature*, 362:175-179 (1993). Briefly, Sf9 cells are washed, resuspended in sonication buffer (25 mL Hepes, pH 7.9; 12.5 mM MgCl₂; 0.1 mM EDTA; 10% Glycerol; 0.1% NP-40; 0.4 M KCl), and sonicated 20 twice for 20 seconds on ice. The sonicates are cleared by centrifugation, and the supernatant is diluted 50-fold in loading buffer (50 mM phosphate, 300 mM NaCl, 10% Glycerol, pH 7.8) and filtered through a 0.45 μm filter. A Ni²⁺-NTA agarose column (commercially available from Qiagen) is prepared with a bed volume of 5 mL, washed with 25 mL of water and equilibrated with 25 mL of loading buffer. The filtered cell extract is loaded onto the column at 0.5 mL per minute. The column is washed to baseline A₂₈₀ with loading buffer, at 25 which point fraction collection is started. Next, the column is washed with a secondary wash buffer (50 mM phosphate; 300 mM NaCl, 10% Glycerol, pH 6.0), which elutes nonspecifically bound protein. After reaching A₂₈₀ baseline again, the column is developed with a 0 to 500 mM Imidazole gradient in the secondary wash buffer. One mL fractions are collected and analyzed by SDS-PAGE and silver staining or western blot with Ni²⁺-NTA-conjugated to alkaline phosphatase (Qiagen). Fractions containing the eluted His₁₀-tagged PRO 30 polypeptide are pooled and dialyzed against loading buffer.

Alternatively, purification of the IgG tagged (or Fc tagged) PRO polypeptide can be performed using known chromatography techniques, including for instance, Protein A or protein G column chromatography.

PRO211, PRO217, PRO230, PRO187, PRO265, PRO246, PRO228, PRO533, PRO245, PRO221, PRO220, PRO258, PRO266, PRO269, PRO287, PRO214, PRO301, PRO224, PRO222, PRO234, PRO231, 35 PRO229, PRO235, PRO239, PRO257, PRO272, PRO294, PRO295, PRO328, PRO326, PRO331, PRO334, PRO346 and PRO310 were successfully expressed in baculovirus infected Sf9 or high5 insect cells. While the expression was actually performed in a 0.5-2 L scale, it can be readily scaled up for larger (e.g. 8 L)

preparations. The proteins were expressed as an IgG construct (immunoadhesin), in which the protein extracellular region was fused to an IgG1 constant region sequence containing the hinge, CH2 and CH3 domains and/or in poly-His tagged forms.

Following PCR amplification, the respective coding sequences were subcloned into a baculovirus expression vector (pb.PH.IgG for IgG fusions and pb.PH.His.c for poly-His tagged proteins), and the vector and Baculogold® baculovirus DNA (Pharmingen) were co-transfected into 105 *Spodoptera frugiperda* ("Sf9") cells (ATCC CRL 1711), using Lipofectin (Gibco BRL). pb.PH.IgG and pb.PH.His are modifications of the commercially available baculovirus expression vector pVL1393 (Pharmingen), with modified polylinker regions to include the His or Fc tag sequences. The cells were grown in Hink's TNM-FH medium supplemented with 10% FBS (Hyclone). Cells were incubated for 5 days at 28°C. The supernatant was harvested and subsequently used for the first viral amplification by infecting Sf9 cells in Hink's TNM-FH medium supplemented with 10% FBS at an approximate multiplicity of infection (MOI) of 10. Cells were incubated for 3 days at 28°C. The supernatant was harvested and the expression of the constructs in the baculovirus expression vector was determined by batch binding of 1 ml of supernatant to 25 mL of Ni-NTA beads (QIAGEN) for histidine tagged proteins or Protein-A Sepharose CL-4B beads (Pharmacia) for IgG tagged proteins followed by SDS-PAGE analysis comparing to a known concentration of protein standard by Coomassie blue staining.

The first viral amplification supernatant was used to infect a spinner culture (500 ml) of Sf9 cells grown in ESF-921 medium (Expression Systems LLC) at an approximate MOI of 0.1. Cells were incubated for 3 days at 28°C. The supernatant was harvested and filtered. Batch binding and SDS-PAGE analysis was repeated, as necessary, until expression of the spinner culture was confirmed.

The conditioned medium from the transfected cells (0.5 to 3 L) was harvested by centrifugation to remove the cells and filtered through 0.22 micron filters. For the poly-His tagged constructs, the protein construct were purified using a Ni-NTA column (Qiagen). Before purification, imidazole was added to the conditioned media to a concentration of 5 mM. The conditioned media were pumped onto a 6 ml Ni-NTA column equilibrated in 20 mM Hepes, pH 7.4, buffer containing 0.3 M NaCl and 5 mM imidazole at a flow rate of 4-5 ml/min. at 4°C. After loading, the column was washed with additional equilibration buffer and the protein eluted with equilibration buffer containing 0.25 M imidazole. The highly purified protein was subsequently desalted into a storage buffer containing 10 mM Hepes, 0.14 M NaCl and 4% mannitol, pH 6.8, with a 25 ml G25 Superfine (Pharmacia) column and stored at -80°C.

Immunoadhesin (Fc containing) constructs of proteins were purified from the conditioned media as follows. The conditioned media were pumped onto a 5 ml Protein A column (Pharmacia) which had been equilibrated in 20 mM Na phosphate buffer, pH 6.8. After loading, the column was washed extensively with equilibration buffer before elution with 100 mM citric acid, pH 3.5. The eluted protein was immediately neutralized by collecting 1 ml fractions into tubes containing 275 mL of 1 M Tris buffer, pH 9. The highly purified protein was subsequently desalted into storage buffer as described above for the poly-His tagged proteins. The homogeneity of the proteins was verified by SDS polyacrylamide gel (PEG) electrophoresis and N-terminal amino acid sequencing by Edman degradation.

EXAMPLE 57: Preparation of Antibodies that Bind to PRO Polypeptides

This example illustrates preparation of monoclonal antibodies which can specifically bind to a PRO polypeptide.

Techniques for producing the monoclonal antibodies are known in the art and are described, for instance, in Goding, *supra*. Immunogens that may be employed include purified PRO polypeptide, fusion proteins containing the PRO polypeptide, and cells expressing recombinant PRO polypeptide on the cell surface. Selection of the immunogen can be made by the skilled artisan without undue experimentation.

Mice, such as Balb/c, are immunized with the PRO polypeptide immunogen emulsified in complete Freund's adjuvant and injected subcutaneously or intraperitoneally in an amount from 1-100 micrograms. Alternatively, the immunogen is emulsified in MPL-TDM adjuvant (Ribi Immunochemical Research, Hamilton, MT) and injected into the animal's hind foot pads. The immunized mice are then boosted 10 to 12 days later with additional immunogen emulsified in the selected adjuvant. Thereafter, for several weeks, the mice may also be boosted with additional immunization injections. Serum samples may be periodically obtained from the mice by retro-orbital bleeding for testing in ELISA assays to detect anti-PRO polypeptide antibodies.

After a suitable antibody titer has been detected, the animals "positive" for antibodies can be injected with a final intravenous injection of PRO polypeptide. Three to four days later, the mice are sacrificed and the spleen cells are harvested. The spleen cells are then fused (using 35% polyethylene glycol) to a selected murine myeloma cell line such as P3X63AgU.1, available from ATCC, No. CRL 1597. The fusions generate hybridoma cells which can then be plated in 96 well tissue culture plates containing HAT (hypoxanthine, aminopterin, and thymidine) medium to inhibit proliferation of non-fused cells, myeloma hybrids, and spleen cell hybrids.

The hybridoma cells will be screened in an ELISA for reactivity against the PRO polypeptide. Determination of "positive" hybridoma cells secreting the desired monoclonal antibodies against the PRO polypeptide is within the skill in the art.

The positive hybridoma cells can be injected intraperitoneally into syngeneic Balb/c mice to produce ascites containing the anti-PRO polypeptide monoclonal antibodies. Alternatively, the hybridoma cells can be grown in tissue culture flasks or roller bottles. Purification of the monoclonal antibodies produced in the ascites can be accomplished using ammonium sulfate precipitation, followed by gel exclusion chromatography. Alternatively, affinity chromatography based upon binding of antibody to protein A or protein G can be employed.

30

EXAMPLE 58: Chimeric PRO Polypeptides

PRO polypeptides may be expressed as chimeric proteins with one or more additional polypeptide domains added to facilitate protein purification. Such purification facilitating domains include, but are not limited to, metal chelating peptides such as histidine-tryptophan modules that allow purification on immobilized metals, protein A domains that allow purification on immobilized immunoglobulin, and the domain utilized in the FLAGSTTM extension/affinity purification system (Immunex Corp., Seattle Wash.). The inclusion of a cleavable linker sequence such as Factor XA or enterokinase (Invitrogen, San Diego Calif.) between the

purification domain and the PRO polypeptide sequence may be useful to facilitate expression of DNA encoding the PRO polypeptide.

EXAMPLE 59: Purification of PRO Polypeptides Using Specific Antibodies

5 Native or recombinant PRO polypeptides may be purified by a variety of standard techniques in the art of protein purification. For example, pro-PRO polypeptide, mature PRO polypeptide, or pre-PRO polypeptide is purified by immunoaffinity chromatography using antibodies specific for the PRO polypeptide of interest. In general, an immunoaffinity column is constructed by covalently coupling the anti-PRO polypeptide antibody to an activated chromatographic resin.

10 Polyclonal immunoglobulins are prepared from immune sera either by precipitation with ammonium sulfate or by purification on immobilized Protein A (Pharmacia LKB Biotechnology, Piscataway, N.J.). Likewise, monoclonal antibodies are prepared from mouse ascites fluid by ammonium sulfate precipitation or chromatography on immobilized Protein A. Partially purified immunoglobulin is covalently attached to a chromatographic resin such as CnBr-activated SEPHAROSE™ (Pharmacia LKB Biotechnology). The antibody 15 is coupled to the resin, the resin is blocked, and the derivative resin is washed according to the manufacturer's instructions.

Such an immunoaffinity column is utilized in the purification of PRO polypeptide by preparing a fraction from cells containing PRO polypeptide in a soluble form. This preparation is derived by solubilization of the whole cell or of a subcellular fraction obtained via differential centrifugation by the addition of detergent or by 20 other methods well known in the art. Alternatively, soluble PRO polypeptide containing a signal sequence may be secreted in useful quantity into the medium in which the cells are grown.

A soluble PRO polypeptide-containing preparation is passed over the immunoaffinity column, and the column is washed under conditions that allow the preferential absorbance of PRO polypeptide (e.g., high ionic strength buffers in the presence of detergent). Then, the column is eluted under conditions that disrupt 25 antibody/PRO polypeptide binding (e.g., a low pH buffer such as approximately pH 2-3, or a high concentration of a chaotrope such as urea or thiocyanate ion), and PRO polypeptide is collected.

EXAMPLE 60: Drug Screening

This invention is particularly useful for screening compounds by using PRO polypeptides or binding 30 fragment thereof in any of a variety of drug screening techniques. The PRO polypeptide or fragment employed in such a test may either be free in solution, affixed to a solid support, borne on a cell surface, or located intracellularly. One method of drug screening utilizes eukaryotic or prokaryotic host cells which are stably transformed with recombinant nucleic acids expressing the PRO polypeptide or fragment. Drugs are screened against such transformed cells in competitive binding assays. Such cells, either in viable or fixed form, can be 35 used for standard binding assays. One may measure, for example, the formation of complexes between PRO polypeptide or a fragment and the agent being tested. Alternatively, one can examine the diminution in complex formation between the PRO polypeptide and its target cell or target receptors caused by the agent being tested.

Thus, the present invention provides methods of screening for drugs or any other agents which can affect a PRO polypeptide-associated disease or disorder. These methods comprise contacting such an agent with an PRO polypeptide or fragment thereof and assaying (I) for the presence of a complex between the agent and the PRO polypeptide or fragment, or (ii) for the presence of a complex between the PRO polypeptide or fragment and the cell, by methods well known in the art. In such competitive binding assays, the PRO polypeptide or 5 fragment is typically labeled. After suitable incubation, free PRO polypeptide or fragment is separated from that present in bound form, and the amount of free or uncomplexed label is a measure of the ability of the particular agent to bind to PRO polypeptide or to interfere with the PRO polypeptide/cell complex.

Another technique for drug screening provides high throughput screening for compounds having suitable binding affinity to a polypeptide and is described in detail in WO 84/03564, published on September 13, 1984. 10 Briefly stated, large numbers of different small peptide test compounds are synthesized on a solid substrate, such as plastic pins or some other surface. As applied to a PRO polypeptide, the peptide test compounds are reacted with PRO polypeptide and washed. Bound PRO polypeptide is detected by methods well known in the art. Purified PRO polypeptide can also be coated directly onto plates for use in the aforementioned drug screening techniques. In addition, non-neutralizing antibodies can be used to capture the peptide and immobilize it on the 15 solid support.

This invention also contemplates the use of competitive drug screening assays in which neutralizing antibodies capable of binding PRO polypeptide specifically compete with a test compound for binding to PRO polypeptide or fragments thereof. In this manner, the antibodies can be used to detect the presence of any peptide which shares one or more antigenic determinants with PRO polypeptide.

20

EXAMPLE 61: Rational Drug Design

The goal of rational drug design is to produce structural analogs of biologically active polypeptide of interest (*i.e.*, a PRO polypeptide) or of small molecules with which they interact, *e.g.*, agonists, antagonists, or inhibitors. Any of these examples can be used to fashion drugs which are more active or stable forms of the 25 PRO polypeptide or which enhance or interfere with the function of the PRO polypeptide *in vivo* (*c.f.*, Hodgson, Bio/Technology, 9: 19-21 (1991)).

In one approach, the three-dimensional structure of the PRO polypeptide, or of an PRO polypeptide-inhibitor complex, is determined by x-ray crystallography, by computer modeling or, most typically, by a combination of the two approaches. Both the shape and charges of the PRO polypeptide must be ascertained 30 to elucidate the structure and to determine active site(s) of the molecule. Less often, useful information regarding the structure of the PRO polypeptide may be gained by modeling based on the structure of homologous proteins. In both cases, relevant structural information is used to design analogous PRO polypeptide-like molecules or to identify efficient inhibitors. Useful examples of rational drug design may include molecules which have improved activity or stability as shown by Braxton and Wells, Biochemistry, 31:7796-7801 (1992) or which act as 35 inhibitors, agonists, or antagonists of native peptides as shown by Athauda *et al.*, J. Biochem., 113:742-746 (1993).

It is also possible to isolate a target-specific antibody, selected by functional assay, as described above,

and then to solve its crystal structure. This approach, in principle, yields a pharmacore upon which subsequent drug design can be based. It is possible to bypass protein crystallography altogether by generating anti-idiotypic antibodies (anti-ids) to a functional, pharmacologically active antibody. As a mirror image of a mirror image, the binding site of the anti-ids would be expected to be an analog of the original receptor. The anti-id could then be used to identify and isolate peptides from banks of chemically or biologically produced peptides. The isolated 5 peptides would then act as the pharmacore.

By virtue of the present invention, sufficient amounts of the PRO polypeptide may be made available to perform such analytical studies as X-ray crystallography. In addition, knowledge of the PRO polypeptide amino acid sequence provided herein will provide guidance to those employing computer modeling techniques in place of or in addition to x-ray crystallography.

10

EXAMPLE 62: Diagnostic Test Using PRO317 Polypeptide-Specific Antibodies

Particular anti-PRO317 polypeptide antibodies are useful for the diagnosis of prepathologic conditions, and chronic or acute diseases such as gynecological diseases or ischemic diseases which are characterized by differences in the amount or distribution of PRO317. PRO317 has been found to be expressed in human kidney 15 and is thus likely to be associated with abnormalities or pathologies which affect this organ. Further, since it is so closely related to EBAF-1, it is likely to affect the endometrium and other genital tissues. Further, due to library sources of certain ESTs, it appears that PRO317 may be involved as well in forming blood vessels and hence to be a modulator of angiogenesis.

Diagnostic tests for PRO317 include methods utilizing the antibody and a label to detect PRO317 in 20 human body fluids, tissues, or extracts of such tissues. The polypeptide and antibodies of the present invention may be used with or without modification. Frequently, the polypeptide and antibodies will be labeled by joining them, either covalently or noncovalently, with a substance which provides for a detectable signal. A wide variety of labels and conjugation techniques are known and have been reported extensively in both the scientific and patent literature. Suitable labels include radionuclides, enzymes, substrates, cofactors, inhibitors, fluorescent 25 agents, chemiluminescent agents, magnetic particles, and the like. Patents teaching the use of such labels include U.S. Pat. Nos. 3,817,837; 3,850,752; 3,939,350; 3,996,345; 4,277,437; 4,275,149; and 4,366,241. Also, recombinant immunoglobulins may be produced as shown in U.S. Pat. No. 4,816,567.

A variety of protocols for measuring soluble or membrane-bound PRO317, using either polyclonal or 30 monoclonal antibodies specific for that PRO317, are known in the art. Examples include enzyme-linked immunosorbent assay (ELISA), radioimmunoassay (RIA), radioreceptor assay (RRA), and fluorescent activated cell sorting (FACS). A two-site monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on PRO317 is preferred, but a competitive binding assay may be employed. These assays are described, among other places, in Maddox *et al.* J Exp. Med., 158:1211 (1983).

35 EXAMPLE 63: Identification of PRO317 Receptors

Purified PRO317 is useful for characterization and purification of specific cell surface receptors and other binding molecules. Cells which respond to PRO317 by metabolic changes or other specific responses are

likely to express a receptor for PRO317. Such receptors include, but are not limited to, receptors associated with and activated by tyrosine and serine/threonine kinases. See Kolodziejczyk and Hall, *supra*, for a review on known receptors for the TGF- superfamily. Candidate receptors for this superfamily fall into two primary groups, termed type I and type II receptors. Both types are serine/threonine kinases. Upon activation by the appropriate ligand, type I and type II receptors physically interact to form hetero-oligomers and subsequently 5 activate intracellular signaling cascades, ultimately regulating gene transcription and expression. In addition, TGF- binds to a third receptor class, type III, a membrane-anchored proteoglycan lacking the kinase activity typical of signal transducing molecules.

PRO317 receptors or other PRO317-binding molecules may be identified by interaction with radiolabeled PRO317. Radioactive labels may be incorporated into PRO317 by various methods known in the 10 art. A preferred embodiment is the labeling of primary amino groups in PRO317 with ^{125}I Bolton-Hunter reagent (Bolton and Hunter, *Biochem. J.*, 133:529 (1973)), which has been used to label other polypeptides without concomitant loss of biological activity (Hebert *et al.*, *J. Biol. Chem.*, 266:18989 (1991); McColl *et al.*, *J. Immunol.*, 150:4550-4555 (1993)). Receptor-bearing cells are incubated with labeled PRO317. The cells are then washed to remove unbound PRO317, and receptor-bound PRO317 is quantified. The data obtained using 15 different concentrations of PRO317 are used to calculate values for the number and affinity of receptors.

Labeled PRO317 is useful as a reagent for purification of its specific receptor. In one embodiment of affinity purification, PRO317 is covalently coupled to a chromatography column. Receptor-bearing cells are extracted, and the extract is passed over the column. The receptor binds to the column by virtue of its biological affinity for PRO317. The receptor is recovered from the column and subjected to N-terminal protein sequencing. 20 This amino acid sequence is then used to design degenerate oligonucleotide probes for cloning the receptor gene.

In an alternative method, mRNA is obtained from receptor-bearing cells and made into a cDNA library. The library is transfected into a population of cells, and those cells expressing the receptor are selected using fluorescently labeled PRO317. The receptor is identified by recovering and sequencing recombinant DNA from 25 highly labeled cells.

25 In another alternative method, antibodies are raised against the surface of receptor bearing cells, specifically monoclonal antibodies. The monoclonal antibodies are screened to identify those which inhibit the binding of labeled PRO317. These monoclonal antibodies are then used in affinity purification or expression cloning of the receptor.

Soluble receptors or other soluble binding molecules are identified in a similar manner. Labeled 30 PRO317 is incubated with extracts or other appropriate materials derived from the uterus. After incubation, PRO317 complexes larger than the size of purified PRO317 are identified by a sizing technique such as size-exclusion chromatography or density gradient centrifugation and are purified by methods known in the art. The soluble receptors or binding protein(s) are subjected to N-terminal sequencing to obtain information sufficient for database identification, if the soluble protein is known, or for cloning, if the soluble protein is unknown.

35

EXAMPLE 64: Determination of PRO317-Induced Cellular Response

The biological activity of PRO317 is measured, for example, by binding of an PRO317 of the invention

to an PRO317 receptor. A test compound is screened as an antagonist for its ability to block binding of PRO317 to the receptor. A test compound is screened as an agonist of the PRO317 for its ability to bind an PRO317 receptor and influence the same physiological events as PRO317 using, for example, the KIRA-ELISA assay described by Sadick *et al.*, *Analytical Biochemistry*, 235:207-214 (1996) in which activation of a receptor tyrosine kinase is monitored by immuno-capture of the activated receptor and quantitation of the level of ligand-induced phosphorylation. The assay may be adapted to monitor PRO317-induced receptor activation through the use of an PRO317 receptor-specific antibody to capture the activated receptor. These techniques are also applicable to other PRO polypeptides described herein.

EXAMPLE 65: Use of PRO224 for Screening Compounds

10 PRO224 is expressed in a cell stripped of membrane proteins and capable of expressing PRO224. Low density lipoproteins having a detectable label are added to the cells and incubated for a sufficient time for endocytosis. The cells are washed. The cells are then analysed for label bound to the membrane and within the cell after cell lysis. Detection of the low density lipoproteins within the cell determines that PRO224 is within the family of low density lipoprotein receptor proteins. Members found within this family are then used for 15 screening compounds which affect these receptors, and particularly the uptake of cholesterol via these receptors.

EXAMPLE 66: Ability of PRO Polypeptides to Inhibit Vascular Endothelial Growth Factor (VEGF) Stimulated Proliferation of Endothelial Cell Growth

20 The ability of various PRO polypeptides to inhibit VEGF stimulated proliferation of endothelial cells was tested. Specifically, bovine adrenal cortical capillary endothelial (ACE) cells (from primary culture, maximum 12-14 passages) were plated on 96-well microtiter plates (Amersham Life Science) at a density of 500 cells/well per 100 μ L in low glucose DMEM, 10% calf serum, 2 mM glutamine, 1x pen/strept and fungizone, supplemented with 3 ng/mL VEGF. Controls were plated the same way but some did not include VEGF. A test sample of the PRO polypeptide of interest was added in a 100 μ L volume for a 200 μ L final volume. Cells 25 were incubated for 6-7 days at 37°C. The media was aspirated and the cells washed 1x with PBS. An acid phosphatase reaction mixture (100 μ L, 0.1M sodium acetate, pH 5.5, 0.1% Triton-100, 10 mM p-nitrophenyl phosphate) was added. After incubation for 2 hours at 37°C, the reaction was stopped by addition of 10 μ L 1N NaOH. OD was measured on microtiter plate reader at 405 nm. Controls were no cells, cells alone, cells + FGF (5 ng/mL), cells + VEGF (3 ng/mL), cells + VEGF (3 ng/ml) + TGF- β (1 ng/ml), and cells + VEGF 30 (3ng/mL) + LIF (5 ng/mL). (TGF- β at a 1 ng/ml concentration is known to block 70-90% of VEGF stimulated cell proliferation.)

35 The results were assessed by calculating the percentage inhibition of VEGF (3 ng/ml) stimulated cells proliferation, determined by measuring acid phosphatase activity at OD405 nm, (1) relative to cells without stimulation, and (2) relative to the reference TGF- β inhibition of VEGF stimulated activity. The results, as shown in Table 2 below, are indicative of the utility of the PRO polypeptides in cancer therapy and specifically in inhibiting tumor angiogenesis. The numerical values (relative inhibition) shown in Table 2 are determined by calculating the percent inhibition of VEGF stimulated proliferation by the PRO polypeptide relative to cells

without stimulation and then dividing that percentage into the percent inhibition obtained by TGF- β at 1 ng/ml which is known to block 70-90% of VEGF stimulated cell proliferation.

Table 2

	<u>PRO Name</u>	<u>PRO Concentration</u>	<u>Relative Inhibition</u>
5	PRO211	0.01%	99.0
	PRO211	0.01%	1.09
	PRO211	0.1%	0.95
	PRO211	0.1%	67.0
10	PRO211	1.0%	0.27
	PRO211	1.0%	20.0
	PRO217	0.01%	1.06
	PRO217	0.1%	0.84
	PRO217	1.0%	0.39
15	PRO217	2.5 μ M	0.2
	PRO217	25 nM	0.88
	PRO217	250 nM	0.58
	PRO187	0.01%	0.91
	PRO187	0.1%	0.82
20	PRO187	1.0%	0.44
	PRO219	5.7 μ M	0.61
	PRO219	57 nM	1.09
	PRO219	570 nM	0.97
	PRO246	0.01%	1.04
25	PRO246	0.1%	1.0
	PRO246	1.0%	0.49
	PRO228	0.01%	0.99
	PRO228	0.1%	0.93
	PRO228	1.0%	0.57
30	PRO228	0.01%	0.95
	PRO228	0.01%	0.98
	PRO228	0.1%	0.77
	PRO228	0.1%	0.88
	PRO228	1.0%	0.16
35	PRO228	1.0%	0.48
	PRO245	0.01%	0.76
	PRO245	0.1%	0.35
	PRO245	1.0%	0.11
	PRO245	0.48 nM	1.03
40	PRO245	4.8 nM	0.95
	PRO245	48 nM	0.49
	PRO221	0.01%	1.03
	PRO221	0.01%	1.06
	PRO221	0.1%	0.82
45	PRO221	0.1%	0.93
	PRO221	1.0%	0.31
	PRO221	1.0%	0.43
	PRO258	0.01%	0.98
	PRO258	0.01%	1.06
50	PRO258	0.1%	0.95
	PRO258	0.1%	1.02
	PRO258	1.0%	0.6
	PRO258	1.0%	0.69

PAGE INTENTIONALLY LEFT BLANK

Table 2 cont'

PRO Name	PRO Concentration	Relative Inhibition
PRO301	7.0 μ M	1.02
PRO301	70 μ M	0.88
5 PRO301	700 μ M	0.44
PRO301	0.01 %	0.92
PRO301	0.1 %	0.85
PRO301	1.0 %	0.68
PRO224	0.01 %	101.0
10 PRO224	0.1 %	65.0
PRO224	1.0 %	23.0
PRO272	0.01 %	0.95
PRO272	0.1 %	0.57
PRO272	1.0 %	0.18
15 PRO328	0.01 %	0.98
PRO328	0.1 %	0.96
PRO328	1.0 %	0.6
PRO331	0.01 %	0.88
PRO331	0.1 %	0.82
20 PRO331	1.0 %	0.56

EXAMPLE 67: Retinal Neuron Survival

This example demonstrates that PRO220 polypeptides have efficacy in enhancing the survival of retinal neuron cells.

25 Sprague Dawley rat pups at postnatal day 7 (mixed population: glia and retinal neuronal types) are killed by decapitation following CO₂ anesthesia and the eyes are removed under sterile conditions. The neural retina is dissected away from the pigment epithelium and other ocular tissue and then dissociated into a single cell suspension using 0.25% trypsin in Ca²⁺, Mg²⁺-free PBS. The retinas are incubated at 37°C for 7-10 minutes after which the trypsin is inactivated by adding 1 ml soybean trypsin inhibitor. The cells are plated at 100,000 cells per well in 96 well plates in DMEM/F12 supplemented with N2 and with or without the specific test PRO polypeptide. Cells for all experiments are grown at 37°C in a water saturated atmosphere of 5% CO₂. After 30 2-3 days in culture, cells are stained with calcein AM then fixed using 4% paraformaldehyde and stained with DAPI for determination of total cell count. The total cells (fluorescent) are quantified at 20X objective magnification using CCD camera and NIH image software for MacIntosh. Fields in the well are chosen at random.

35 The effect of various concentration of PRO220 polypeptides are reported in Table 3 below where percent survival is calculated by dividing the total number of calcein AM positive cells at 2-3 days in culture by the total number of DAPI-labeled cells at 2-3 days in culture. Anything above 30% survival is considered positive.

Table 3

<u>PRO Name</u>	<u>PRO Concentration</u>	<u>Percent Survival</u>
PRO220	0.01%	2.4%
PRO220	0.01%	4.1%
PRO220	0.1%	3.0%
5 PRO220	0.1%	3.1%
PRO220	1.0%	72.4%
PRO220	1.0%	42.1%

EXAMPLE 68: Rod Photoreceptor Survival

10 This example demonstrates that PRO220 polypeptides have efficacy in enhancing the survival of rod photoreceptor cells.

Sprague Dawley rat pups at 7 day postnatal (mixed population: glia and retinal neuronal cell types) are killed by decapitation following CO₂ anesthesia and the eyes are removed under sterile conditions. The neural retina is dissected away from the pigment epithelium and other ocular tissue and then dissociated into a single cell suspension using 0.25% trypsin in Ca²⁺, Mg²⁺-free PBS. The retinas are incubated at 37°C for 7-10 minutes after which the trypsin is inactivated by adding 1 ml soybean trypsin inhibitor. The cells are plated at 100,000 cells per well in 96 well plates in DMEM/F12 supplemented with N2 and with or without the specific test PRO polypeptide. Cells for all experiments are grown at 37°C in a water saturated atmosphere of 5% CO₂. After 2-3 days in culture, cells are fixed using 4% paraformaldehyde, and then stained using CellTracker Green CMFDA. Rho 4D2 (ascites or IgG 1:100), a monoclonal antibody directed towards the visual pigment rhodopsin is used to detect rod photoreceptor cells by indirect immunofluorescence. The results are reported as % survival: total number of calcein/CellTracker - rhodopsin positive cells at 2-3 days in culture, divided by the total number of rhodopsin positive cells at time 2-3 days in culture. The total cells (fluorescent) are quantified at 20x objective magnification using a CCD camera and NIH image software for MacIntosh. Fields in the well are chosen at random.

20 The effect of various concentration of PRO220 polypeptides are reported in Table 4 below. Anything above 10% survival is considered positive..

Table 4

	<u>PRO Name</u>	<u>PRO Concentration</u>	<u>Percent Survival</u>
30	PRO220	0.01%	0.0%
	PRO220	0.1%	0.0%
	PRO220	2.0%	0.0%
	PRO220	10%	0.0%
35	PRO220	20%	66.9%
	PRO220	1.0%	56.9%

EXAMPLE 69: Induction of Endothelial Cell Apoptosis

40 The ability of PRO228 polypeptides to induce apoptosis in endothelial cells was tested in human venous umbilical vein endothelial cells (HUVEC, Cell Systems), using a 96-well format, in 0% serum media supplemented with 100 ng/ml VEGF. (As HUVEC cells are easily dislodged from the plating surface, all pipetting in the wells must be done as gently as practicable.)

The media was aspirated and the cells washed once with PBS. 5 ml of 1 x trypsin was added to the cells in a T-175 flask, and the cells were allowed to stand until they were released from the plate (about 5-10 minutes). Trypsinization was stopped by adding 5 ml of growth media. The cells were spun at 1000 rpm for 5 minutes at 4°C. The media was aspirated and the cells were resuspended in 10 ml of 10% serum complemented medium (Cell Systems), 1 x penn/strep.

5 The cells were plated on 96-well microtiter plates (Amersham Life Science, cytostar-T scintillating microplate, RPNQ160, sterile, tissue-culture treated, individually wrapped), in 10% serum (CSG-medium, Cell Systems), at a density of 2×10^4 cells per well in a total volume of 100 μ l. The PRO228 polypeptide was added in triplicate at dilutions of 1%, 0.33% and 0.11%. Wells without cells were used as a blank and wells with cells only as a negative control. As a positive control 1:3 serial dilutions of 50 μ l of a 3x stock of staurosporine were
10 used. The ability of the PRO228 polypeptide to induce apoptosis was determined using Annexin V, a member of the calcium and phospholipid binding proteins, to detect apoptosis.

15 0.2 ml Annexin V - Biotin stock solution (100 μ g/ml) were diluted in 4.6 ml 2 x Ca^{2+} binding buffer and 2.5% BSA (1:25 dilution). 50 μ ls of the diluted Annexin V - Biotin solution were added to each well (except controls) to a final concentration of 1.0 μ g/ml. The samples were incubated for 10-15 minutes with Annexin-
20 Biotin prior to direct addition of ^{35}S -Streptavidin. ^{35}S -Streptavidin was diluted in 2x Ca^{2+} binding buffer, 2.5% BSA and was added to all wells at a final concentration of 3×10^4 cpm/well. The plates were then sealed, centrifuged at 1000 rpm for 15 minutes and placed on orbital shaker for 2 hours. The analysis was performed on 1450 Microbeta Trilux (Wallac). The results are shown in Table 5 below where percent above background represents the percentage amount of counts per minute above the negative controls. Percents greater than or equal to 30% above background are considered positive.

Table 5

<u>PRO Name</u>	<u>PRO Concentration</u>	<u>Percent Above Background</u>
PRO228	0.11%	0.7%
PRO228	0.11%	47.6%
PRO228	0.33%	92.2%
PRO228	0.33%	123.7%
PRO228	1.0%	51.4%
PRO228	1.0%	95.3%

30

EXAMPLE 70: PDB12 Cell Inhibition

This example demonstrates that various PRO polypeptides have efficacy in inhibiting protein production by PDB12 pancreatic ductal cells.

35 PDB12 pancreatic ductal cells are plated on fibronectin coated 96 well plates at 1.5×10^3 cells per well in 100 μ L/180 μ L of growth media. 100 μ L of growth media with the PRO polypeptide test sample or negative control lacking the PRO polypeptide is then added to well, for a final volume of 200 μ L. Controls contain growth medium containing a protein shown to be inactive in this assay. Cells are incubated for 4 days at 37°C. 20 μ L of Alamar Blue Dye (AB) is then added to each well and the fluorescent reading is measured at 4 hours post addition of AB, on a microtiter plate reader at 530 nm excitation and 590 nm emission. The standard employed
40 is cells without Bovine Pituitary Extract (BPE) and with various concentrations of BPE. Buffer or CM controls

from unknowns are run 2 times on each 96 well plate.

The results from these assays are shown in Table 6 below wherein percent decrease in protein production is calculated by comparing the Alamar Blue Dye calculated protein concentration produced by the PRO polypeptide-treated cells with the Alamar Blue Dye calculated protein concentration produced by the negative control cells. A percent decrease in protein production of greater than or equal to 25% as compared to the negative control cells is considered positive.

5

Table 6

	<u>PRO Name</u>	<u>PRO Concentration</u>	<u>Percent Decrease in Protein Production</u>
10	PRO211	0.1%	0.0%
	PRO211	0.01%	0.6%
	PRO211	1.0%	59.7%
	PRO287	2.0%	22.3%
	PRO287	10%	18.2%
	PRO287	50%	67.5%
15	PRO287	2.0%	45.53%
	PRO287	10%	57.3%
	PRO287	50%	52.24%
	PRO301	2.0%	0.0%
	PRO301	10%	59.8%
20	PRO301	50%	65.6%
	PRO293	2.0%	0.0%
	PRO293	10%	40.4%
	PRO293	50%	56.7%

25 EXAMPLE 71: Stimulation of Adult Heart Hypertrophy

This assay is designed to measure the ability of various PRO polypeptides to stimulate hypertrophy of adult heart.

Ventricular myocytes freshly isolated from adult (250g) Sprague Dawley rats are plated at 2000 cell/well in 180 μ l volume. Cells are isolated and plated on day 1, the PRO polypeptide-containing test samples or growth medium only (negative control) (20 μ l volume) is added on day 2 and the cells are then fixed and stained on day 5. After staining, cell size is visualized wherein cells showing no growth enhancement as compared to control cells are given a value of 0.0, cells showing small to moderate growth enhancement as compared to control cells are given a value of 1.0 and cells showing large growth enhancement as compared to control cells are given a value of 2.0. Any degree of growth enhancement as compared to the negative control cells is considered positive

35 for the assay. The results are shown in Table 7 below.

Table 7

	<u>PRO Name</u>	<u>PRO Concentration</u>	<u>Growth Enhancement Score</u>
40	PRO287	20%	1.0
	PRO287	20%	1.0
	PRO301	20%	1.0
	PRO301	20%	1.0
	PRO293	20%	1.0
	PRO293	20%	1.0
	PRO303	20%	1.0
45	PRO303	20%	1.0

EXAMPLE 72: PDB12 Cell Proliferation

This example demonstrates that various PRO polypeptides have efficacy in inducing proliferation of PDB12 pancreatic ductal cells.

PDB12 pancreatic ductal cells are plated on fibronectin coated 96 well plates at 1.5×10^3 cells per well in $100 \mu\text{L}/180 \mu\text{L}$ of growth media. $100 \mu\text{L}$ of growth media with the PRO polypeptide test sample or negative control lacking the PRO polypeptide is then added to well, for a final volume of $200 \mu\text{L}$. Controls contain growth medium containing a protein shown to be inactive in this assay. Cells are incubated for 4 days at 37°C . $20 \mu\text{L}$ of Alamar Blue Dye (AB) is then added to each well and the fluorescent reading is measured at 4 hours post addition of AB, on a microtiter plate reader at 530 nm excitation and 590 nm emission. The standard employed is cells without Bovine Pituitary Extract (BPE) and with various concentrations of BPE. Buffer or growth medium only controls from unknowns are run 2 times on each 96 well plate.

The results from these assays are shown in Table 8 below wherein percent increase in protein production is calculated by comparing the Alamar Blue Dye calculated protein concentration produced by the PRO polypeptide-treated cells with the Alamar Blue Dye calculated protein concentration produced by the negative control cells. A percent increase in protein production of greater than or equal to 25% as compared to the negative control cells is considered positive.

Table 8

<u>PRO Name</u>	<u>PRO Concentration</u>	<u>Percent Increase in Protein Production</u>
PRO301	2.0%	44.0%
PRO301	10%	67.4%
PRO301	50%	185.8%
PRO303	2.0%	27.9%
PRO303	10%	174.9%
PRO303	50%	193.1%

EXAMPLE 73: Enhancement of Heart Neonatal Hypertrophy Induced by PRO224

This assay is designed to measure the ability of PRO224 polypeptides to stimulate hypertrophy of neonatal heart.

Cardiac myocytes from 1-day old Harlan Sprague Dawley rats were obtained. Cells ($180 \mu\text{l}$ at $7.5 \times 10^4/\text{ml}$, serum $<0.1\%$, freshly isolated) are added on day 1 to 96-well plates previously coated with DMEM/F12 + 4% FCS. Test samples containing the test PRO224 polypeptide or growth medium only (negative control) ($20 \mu\text{l}/\text{well}$) are added directly to the wells on day 1. PGF ($20 \mu\text{l}/\text{well}$) is then added on day 2 at final concentration of 10^{-6} M. The cells are then stained on day 4 and visually scored on day 5, wherein cells showing no increase in size as compared to negative controls are scored 0.0, cells showing a small to moderate increase in size as compared to negative controls are scored 1.0 and cells showing a large increase in size as compared to negative controls are scored 2.0. The results are shown in Table 9 below.

Table 9

PRO Name	PRO Concentration	Growth Enhancement Score
PRO224	0.01%	0.0
PRO224	0.1%	0.0
PRO224	1.0%	1.0

5

EXAMPLE 74: *In situ* Hybridization

In situ hybridization is a powerful and versatile technique for the detection and localization of nucleic acid sequences within cell or tissue preparations. It may be useful, for example, to identify sites of gene expression, analyze the tissue distribution of transcription, identify and localize viral infection, follow changes 10 in specific mRNA synthesis and aid in chromosome mapping.

In situ hybridization was performed following an optimized version of the protocol by Lu and Gillett, Cell Vision 1:169-176 (1994), using PCR-generated ^{33}P -labeled riboprobes. Briefly, formalin-fixed, paraffin-embedded human tissues were sectioned, deparaffinized, deproteinized in proteinase K (20 g/ml) for 15 minutes at 37°C, and further processed for *in situ* hybridization as described by Lu and Gillett, *supra*. A $[^{33}\text{P}]$ UTP-labeled antisense riboprobe was generated from a PCR product and hybridized at 55°C overnight. The slides 15 were dipped in Kodak NTB2 nuclear track emulsion and exposed for 4 weeks.

 ^{33}P -Riboprobe synthesis

6.0 μl (125 mCi) of ^{33}P -UTP (Amersham BF 1002, SA < 2000 Ci/mmol) were speed vac dried. To each tube containing dried ^{33}P -UTP, the following ingredients were added:

20 2.0 μl 5x transcription buffer
 1.0 μl DTT (100 mM)
 2.0 μl NTP mix (2.5 mM : 10 μl ; each of 10 mM GTP, CTP & ATP + 10 μl H₂O)
 1.0 μl UTP (50 μM)
 1.0 μl Rnasin
 25 1.0 μl DNA template (1 μg)
 1.0 μl H₂O
 1.0 μl RNA polymerase (for PCR products T3 = AS, T7 = S, usually)

The tubes were incubated at 37°C for one hour. 1.0 μl RQ1 DNase were added, followed by incubation at 37°C for 15 minutes. 90 μl TE (10 mM Tris pH 7.6/1mM EDTA pH 8.0) were added, and the mixture was 30 pipetted onto DE81 paper. The remaining solution was loaded in a Microcon-50 ultrafiltration unit, and spun using program 10 (6 minutes). The filtration unit was inverted over a second tube and spun using program 2 (3 minutes). After the final recovery spin, 100 μl TE were added. 1 μl of the final product was pipetted on DE81 paper and counted in 6 ml of Biofluor II.

The probe was run on a TBE/urea gel. 1-3 μl of the probe or 5 μl of RNA Mrk III were added to 35 μl of loading buffer. After heating on a 95°C heat block for three minutes, the gel was immediately placed on ice. The wells of gel were flushed, the sample loaded, and run at 180-250 volts for 45 minutes. The gel was wrapped in saran wrap and exposed to XAR film with an intensifying screen in -70°C freezer one hour to overnight.

³⁵P-HybridizationA. Pretreatment of frozen sections

The slides were removed from the freezer, placed on aluminium trays and thawed at room temperature for 5 minutes. The trays were placed in 55°C incubator for five minutes to reduce condensation. The slides were fixed for 10 minutes in 4% paraformaldehyde on ice in the fume hood, and washed in 0.5 x SSC for 5 minutes, at room temperature (25 ml 20 x SSC + 975 ml SQ H₂O). After deproteinization in 0.5 µg/ml proteinase K for 10 minutes at 37°C (12.5 µl of 10 mg/ml stock in 250 ml prewarmed RNase-free RNase buffer), the sections were washed in 0.5 x SSC for 10 minutes at room temperature. The sections were dehydrated in 70%, 95%, 100% ethanol, 2 minutes each.

B. Pretreatment of paraffin-embedded sections

10 The slides were deparaffinized, placed in SQ H₂O, and rinsed twice in 2 x SSC at room temperature, for 5 minutes each time. The sections were deproteinized in 20 µg/ml proteinase K (500 µl of 10 mg/ml in 250 ml RNase-free RNase buffer; 37°C, 15 minutes) - human embryo, or 8 x proteinase K (100 µl in 250 ml RNase buffer, 37°C, 30 minutes) - formalin tissues. Subsequent rinsing in 0.5 x SSC and dehydration were performed as described above.

15 C. Prehybridization

The slides were laid out in a plastic box lined with Box buffer (4 x SSC, 50% formamide) - saturated filter paper. The tissue was covered with 50 µl of hybridization buffer (3.75g Dextran Sulfate + 6 ml SQ H₂O), vortexed and heated in the microwave for 2 minutes with the cap loosened. After cooling on ice, 18.75 ml formamide, 3.75 ml 20 x SSC and 9 ml SQ H₂O were added, the tissue was vortexed well, and incubated at 20 42°C for 1-4 hours.

D. Hybridization

1.0 x 10⁶ cpm probe and 1.0 µl tRNA (50 mg/ml stock) per slide were heated at 95°C for 3 minutes. The slides were cooled on ice, and 48 µl hybridization buffer were added per slide. After vortexing, 50 µl ³⁵P mix were added to 50 µl prehybridization on slide. The slides were incubated overnight at 55°C.

25 E. Washes

Washing was done 2 x 10 minutes with 2xSSC, EDTA at room temperature (400 ml 20 x SSC + 16 ml 0.25M EDTA, V_f=4L), followed by RNaseA treatment at 37°C for 30 minutes (500 µl of 10 mg/ml in 250 ml RNase buffer = 20 µg/ml). The slides were washed 2 x 10 minutes with 2 x SSC, EDTA at room temperature. The stringency wash conditions were as follows: 2 hours at 55°C, 0.1 x SSC, EDTA (20 ml 20 x SSC + 16 ml EDTA, V_f=4L).

F. Oligonucleotides

In situ analysis was performed on a variety of DNA sequences disclosed herein. The oligonucleotides employed for these analyses are as follows.

(1) DNA33094-1131 (PRO217)

35 p1 5'-GGATTCTAATACGACTCACTATAGGGCTCAGAAAAGCGAACAGAGAA-3' (SEQ ID NO:348)

p2 5'-CTATGAAATTAAACCTCACTAAAGGGATGTCTTCCATGCCAACCTTC-3'(SEQ ID NO:349)

(2) DNA33223-1136 (PRO230)

p1 5'-GGATTCTAATACGACTCACTATAGGGCGCGATGTCCACTGGGCTAC-3' (SEQ ID NO:350)

p2 5'-CTATGAAATTAACCCCTCACTAAAGGGACGAGGAAGATGGCGGATGGT-3' (SEQ ID NO:351)

5

(3) DNA34435-1140 (PRO232)

p1 5'-GGATTCTAATACGACTCACTATAGGGACCCACCGCGTCCGGCTGCTT-3'(SEQ ID NO:352)

p2 5'-CTATGAAATTAACCCCTCACTAAAGGGACGGGGACACCACGGACCAGA-3' (SEQ ID NO:353)

10

(4) DNA35639-1172 (PRO246)

p1 5'-GGATTCTAATACGACTCACTATAGGGCTTGCTGCGGTTTTGTTCTG-3'(SEQ ID NO:354)

p2 5'-CTATGAAATTAACCCCTCACTAAAGGGAGCTGCCATCCCCTGGTATT-3' (SEQ ID NO:355)

15

(5) DNA49435-1219 (PRO533)

p1 5'-GGATTCTAATACGACTCACTATAGGGCGGATCCTGGCCGGCCTCTG-3' (SEQ ID NO:356)

p2 5'-CTATGAAATTAACCCCTCACTAAAGGGAGCCGGCATGGTCTCAGTTA-3' (SEQ ID NO:357)

20

(6) DNA35638-1141 (PRO245)

p1 5'-GGATTCTAATACGACTCACTATAGGGCGGAAGATGGCGAGGAGGAG-3'(SEQ ID NO:358)

p2 5'-CTATGAAATTAACCCCTCACTAAAGGGACCAAGGCCACAAACGGAAATC-3' (SEQ ID NO:359)

25

(7) DNA33089-1132 (PRO221)

p1 5'-GGATTCTAATACGACTCACTATAGGGCTGTGCTTCATTCTGCCAGTA-3'(SEQ ID NO:360)

p2 5'-CTATGAAATTAACCCCTCACTAAAGGGAGGGTACAATTAAGGGGTGGAT-3' (SEQ ID NO:361)

30

(8) DNA35918-1174 (PRO258)

p1 5'-GGATTCTAATACGACTCACTATAGGGCCCGCCTCGCTCCTGCTCCTG-3'(SEQ ID NO:362)

p2 5'-CTATGAAATTAACCCCTCACTAAAGGGAGGATTGCCGACCCCTCACAG-3' (SEQ ID NO:363)

35

(9) DNA32286-1191 (PRO214)

p1 5'-GGATTCTAATACGACTCACTATAGGGCCCCTCCTGCCCTCCCTGTCC-3'(SEQ ID NO:364)

p2 5'-CTATGAAATTAACCCTCACTAAAGGGAGTGGTGGCGCGATTATCTGC-3' (SEQ ID NO:365)

(10) DNA33221-1133 (PRO224)

p1 5'-GGATTCTAATACGACTCACTATAGGGCGCAGCGATGGCAGCGATGAGG-3' (SEQ ID

5 NO:366)

p2 5'-CTATGAAATTAACCCTCACTAAAGGGACAGACAGGGCAGAGGGAGTG-3'(SEQ ID NO:367)

(11) DNA35557-1137 (PRO234)

p1 5'-GGATTCTAATACGACTCACTATAGGCCAGGAGGCGTGAGGAGAAC-3'(SEQ ID NO:368)

10 p2 5'-CTATGAAATTAACCCTCACTAAAGGGAAAGACATGTCATCGGGAGTGG-3' (SEQ ID NO:369)

(12) DNA33100-1159 (PRO229)

p1 5'-GGATTCTAATACGACTCACTATAGGCCGGGTGGAGGTGGAACAGAAA-3' (SEQ ID

15 NO:370)

p2 5'-CTATGAAATTAACCCTCACTAAAGGGACACAGACAGAGCCCCATACGC-3' (SEQ ID NO:371)

(13) DNA34431-1177 (PRO263)

20 p1 5'-GGATTCTAATACGACTCACTATAGGCCAGGAAATCCGGATGTCTC-3(SEQ ID NO:372)

p2 5'-CTATGAAATTAACCCTCACTAAAGGGAGTAAGGGATGCCACCGAGTA-3' (SEQ ID NO:373)

(14) DNA38268-1188 (PRO295)

25 p1 5'-GGATTCTAATACGACTCACTATAGGCCAGCTACCCGCAGGAGGAGG-3'(SEQ ID NO:374)

p2 5'-CTATGAAATTAACCCTCACTAAAGGGATCCCAGGTGATGAGGTCCAGA-3' (SEQ ID NO:375)

G. Results

30 *In situ* analysis was performed on a variety of DNA sequences disclosed herein. The results from these analyses are as follows.

(1) DNA33094-1131 (PRO217)

Highly distinctive expression pattern, that does not indicate an obvious biological function. In the human embryo it was expressed in outer smooth muscle layer of the GI tract, respiratory cartilage, branching respiratory epithelium, osteoblasts, tendons, gonad, in the optic nerve head and developing dermis. In the adult expression was observed in the epidermal pegs of the chimp tongue, the basal epithelial/myoepithelial cells of the prostate and urinary bladder. Also expressed in the alveolar lining cells of the adult lung, mesenchymal cells juxtaposed

to erectile tissue in the penis and the cerebral cortex (probably glial cells). In the kidney, expression was only seen in disease, in cells surrounding thyroidized renal tubules.

Human fetal tissues examined (E12-E16 weeks) include: Placenta, umbilical cord, liver, kidney, adrenals, thyroid, lungs, heart, great vessels, oesophagus, stomach, small intestine, spleen, thymus, pancreas, brain, eye, spinal cord, body wall, pelvis and lower limb.

5 Adult human tissues examined: Kidney (normal and end-stage), adrenal, myocardium, aorta, spleen, lymph node, gall bladder, pancreas, lung, skin, eye (inc. retina), prostate, bladder, liver (normal, cirrhotic, acute failure).

Non-human primate tissues examined:

10 (a) Chimp Tissues: Salivary gland, stomach, thyroid, parathyroid, skin, thymus, ovary, lymph node.

(b) Rhesus Monkey Tissues: Cerebral cortex, hippocampus, cerebellum, penis.

(2) DNA33223-1136 (PRO230)

15 Sections show an intense signal associated with arterial and venous vessels in the fetus. In arteries the signal appeared to be confined to smooth-muscle/pericytic cells. The signal is also seen in capillary vessels and in glomeruli. It is not clear whether or not endothelial cells are expressing this mRNA. Expression is also observed in epithelial cells in the fetal lens. Strong expression was also seen in cells within placental trophoblastic villi, these cells lie between the trophoblast and the fibroblast-like cells that express HGF - uncertain histogenesis. In the adult, there was no evidence of expression and the wall of the aorta and most 20 vessels appear to be negative. However, expression was seen over vascular channels in the normal prostate and in the epithelium lining the gallbladder. Insurers expression was seen in the vessels of the soft-tissue sarcoma and a renal cell carcinoma. In summary, this is a molecule that shows relatively specific vascular expression in the fetus as well as in some adult organs. Expression was also observed in the fetal lens and the adult gallbladder.

25 In a secondary screen, vascular expression was observed, similar to that observed above, seen in fetal blocks. Expression is on vascular smooth muscle, rather than endothelium. Expression also seen in smooth muscle of the developing oesophagus, so as reported previously, this molecule is not vascular specific. Expression was examined in 4 lung and 4 breast carcinomas. Substantial expression was seen in vascular smooth muscle of at least 3/4 lung cancers and 2/4 breast cancers. In addition, in one breast carcinoma, expression was 30 observed in peritumoral stromal cells of uncertain histogenesis (possibly myofibroblasts). No endothelial cell expression was observed in this study.

(3) DNA34435-1140 (PRO232)

35 Strong expression in prostatic epithelium and bladder epithelium, lower level of expression in bronchial epithelium. High background / low level expression seen in a number of sites, including among others, bone, blood, chondrosarcoma, adult heart and fetal liver. It is felt that this level of signal represents background, partly because signal at this level was seen over the blood. All other tissues negative.

Human fetal tissues examined (E12-E16 weeks) include: Placenta, umbilical cord, liver, kidney, adrenals, thyroid, lungs, heart, great vessels, oesophagus, stomach, small intestine, spleen, thymus, pancreas, brain, eye, spinal cord, body wall, pelvis, testis and lower limb.

Adult human tissues examined: Kidney (normal and end-stage), adrenal, spleen, lymph node, pancreas, lung, eye (inc. retina), bladder, liver (normal, cirrhotic, acute failure).

5 Non-human primate tissues examined:

Chimp Tissues: adrenal

Rhesus Monkey Tissues: Cerebral cortex, hippocampus

In a secondary screen, expression was observed in the epithelium of the prostate, the superficial layers of the urethelium of the urinary bladder, the urethelium lining the renal pelvis and the urethelium of the ureter (1 out of 2 experiments). The urethra of a rhesus monkey was negative; it is unclear whether this represents a true lack of expression by the urethra, or if it is the result of a failure of the probe to cross react with rhesus tissue. The findings in the prostate and bladder are similar to those previously described using an isotopic detection technique. Expression of the mRNA for this antigen is NOT prostate epithelial specific. The antigen may serve as a useful marker for urethelial derived tissues. Expression in the superficial, post-mitotic cells, of the urinary tract epithelium also suggest that it is unlikely to represent a specific stem cell marker, as this would be expected to be expressed specifically in basal epithelium.

(4) DNA35639-1172 (PRO246)

Strongly expressed in fetal vascular endothelium, including tissues of the CNS. Lower level of expression in adult vasculature, including the CNS. Not obviously expressed at higher levels in tumor vascular endothelium. Signal also seen over bone matrix and adult spleen, not obviously cell associated, probably related to non-specific background at these sites.

Human fetal tissues examined (E12-E16 weeks) include: Placenta, umbilical cord, liver, kidney, adrenals, thyroid, lungs, heart, great vessels, oesophagus, stomach, small intestine, spleen, thymus, pancreas, brain, eye, spinal cord, body wall, pelvis, testis and lower limb.

Adult human tissues examined: Kidney (normal and end-stage), adrenal, spleen, lymph node, pancreas, lung, eye (inc. retina), bladder, liver (normal, cirrhotic, acute failure).

Non-human primate tissues examined:

Chimp Tissues: adrenal

30 Rhesus Monkey Tissues: Cerebral cortex, hippocampus

(5) DNA49435-1219 (PRO533)

Moderate expression over cortical neurones in the fetal brain. Expression over the inner aspect of the fetal retina, possible expression in the developing lens. Expression over fetal skin, cartilage, small intestine, 35 placental villi and umbilical cord. In adult tissues there is an extremely high level of expression over the gallbladder epithelium. Moderate expression over the adult kidney, gastric and colonic epithelia. Low-level expression was observed over many cell types in many tissues, this may be related to stickiness of the probe,

these data should therefore be interpreted with a degree of caution.

Human fetal tissues examined (E12-E16 weeks) include: Placenta, umbilical cord, liver, kidney, adrenals, thyroid, lungs, heart, great vessels, oesophagus, stomach, small intestine, spleen, thymus, pancreas, brain, eye, spinal cord, body wall, pelvis, testis and lower limb.

5 Adult human tissues examined: Kidney (normal and end-stage), adrenal, spleen, lymph node, pancreas, lung, eye (inc. retina), bladder, liver (normal, cirrhotic, acute failure).

Non-human primate tissues examined:

Chimp Tissues: adrenal

Rhesus Monkey Tissues: Cerebral cortex, hippocampus, cerebellum.

10 (6) DNA35638-1141 (PRO245)

Expression observed in the endothelium lining a subset of fetal and placental vessels. Endothelial expression was confined to these tissue blocks. Expression also observed over intermediate trophoblast cells of placenta. All other tissues negative.

Fetal tissues examined (E12-E16 weeks) include: Placenta, umbilical cord, liver, kidney, adrenals, thyroid, lungs, heart, great vessels, oesophagus, stomach, small intestine, spleen, thymus, pancreas, brain, eye, spinal cord, body wall, pelvis and lower limb.

15 Adult tissues examined: Liver, kidney, adrenal, myocardium, aorta, spleen, lymph node, pancreas, lung, skin, cerebral cortex (rm), hippocampus(rm), cerebellum(rm), penis, eye, bladder, stomach, gastric carcinoma, colon, colonic carcinoma, thyroid (chimp), parathyroid (chimp) ovary (chimp) and 20 chondrosarcoma. Acetominophen induced liver injury and hepatic cirrhosis

(7) DNA33089-1132 (PRO221)

Specific expression over fetal cerebral white and grey matter, as well as over neurones in the spinal 25 cord. Probe appears to cross react with rat. Low level of expression over cerebellar neurones in adult rhesus brain. All other tissues negative.

Fetal tissues examined (E12-E16 weeks) include: Placenta, umbilical cord, liver, kidney, adrenals, thyroid, lungs, heart, great vessels, oesophagus, stomach, small intestine, spleen, thymus, pancreas, brain, eye, spinal cord, body wall, pelvis and lower limb.

30 Adult tissues examined: Liver, kidney, adrenal, myocardium, aorta, spleen, lymph node, pancreas, lung, skin, cerebral cortex (rm), hippocampus(rm), cerebellum(rm), penis, eye, bladder, stomach, gastric carcinoma, colon, colonic carcinoma and chondrosarcoma. Acetominophen induced liver injury and hepatic cirrhosis

35 (8) DNA35918-1174 (PRO258)

Strong expression in the nervous system. In the rhesus monkey brain expression is observed in cortical, hippocampal and cerebellar neurones. Expression over spinal neurones in the fetal spinal cord, the developing

brain and the inner aspects of the fetal retina. Expression over developing dorsal root and autonomic ganglia as well as enteric nerves. Expression observed over ganglion cells in the adult prostate. In the rat, there is strong expression over the developing hind brain and spinal cord. Strong expression over interstitial cells in the placental villi. All other tissues were negative.

5 Fetal tissues examined (E12-E16 weeks) include: Placenta, umbilical cord, liver, kidney, adrenals, thyroid, lungs, heart, great vessels, oesophagus, stomach, small intestine, spleen, thymus, pancreas, brain, eye, spinal cord, body wall, pelvis and lower limb.

10 Adult tissues examined: Liver, kidney, renal cell carcinoma, adrenal, aorta, spleen, lymph node, pancreas, lung, myocardium, skin, cerebral cortex (rm), hippocampus(rm), cerebellum(rm), bladder, prostate, stomach, gastric carcinoma, colon, colonic carcinoma, thyroid (chimp), parathyroid (chimp) ovary (chimp) and chondrosarcoma. Acetominophen induced liver injury and hepatic cirrhosis.

(9) DNA32286-1191 (PRO214)

Fetal tissue: Low level throughout mesenchyme. Moderate expression in placental stromal cells in membranous tissues and in thyroid. Low level expression in cortical neurones. Adult tissue: all negative.

15 Fetal tissues examined (E12-E16 weeks) include: Placenta, umbilical cord, liver, kidney, adrenals, thyroid, lungs, heart, great vessels, oesophagus, stomach, small intestine, spleen, thymus, pancreas, brain, eye, spinal cord, body wall, pelvis and lower limb.

Adult tissues examined include: Liver, kidney, adrenal, myocardium, aorta, spleen, lymph node, pancreas, lung and skin.

20

(10) DNA33221-1133 (PRO224)

Expression limited to vascular endothelium in fetal spleen, adult spleen, fetal liver, adult thyroid and adult lymph node (chimp). Additional site of expression is the developing spinal ganglia. All other tissues negative.

25 Human fetal tissues examined (E12-E16 weeks) include: Placenta, umbilical cord, liver, kidney, adrenals, thyroid, lungs, heart, great vessels, oesophagus, stomach, small intestine, spleen, thymus, pancreas, brain, eye, spinal cord, body wall, pelvis and lower limb.

Adult human tissues examined: Kidney (normal and end-stage), adrenal, myocardium, aorta, spleen, lymph node, pancreas, lung, skin, eye (inc. retina), bladder, liver (normal, cirrhotic, acute failure).

30 Non-human primate tissues examined:

Chimp Tissues: Salivary gland, stomach, thyroid, parathyroid, skin, thymus, ovary, lymph node.

Rhesus Monkey Tissues: Cerebral cortex, hippocampus, cerebellum, penis.

(11) DNA35557-1137 (PRO234)

35 Specific expression over developing motor neurones in ventral aspect of the fetal spinal cord (will develop into ventral horns of spinal cord). All other tissues negative. Possible role in growth, differentiation and/or development of spinal motor neurons.

Fetal tissues examined (E12-E16 weeks) include: Placenta, umbilical cord, liver, kidney, adrenals, thyroid, lungs, heart, great vessels, oesophagus, stomach, small intestine, spleen, thymus, pancreas, brain, eye, spinal cord, body wall, pelvis and lower limb.

Adult tissues examined: Liver, kidney, adrenal, myocardium, aorta, spleen, lymph node, pancreas, lung, skin, cerebral cortex (rm), hippocampus(rm), cerebellum(rm), penis, eye, bladder, stomach, 5 gastric carcinoma, colon, colonic carcinoma and chondrosarcoma. Acetominophen induced liver injury and hepatic cirrhosis

(12) DNA33100-1159 (PRO229)

Striking expression in mononuclear phagocytes (macrophages) of fetal and adult spleen, liver, lymph 10 node and adult thymus (in tingible body macrophages). The highest expression is in the spleen. All other tissues negative. Localisation and homology are entirely consistent with a role as a scavenger receptor for cells of the reticuloendothelial system. Expression also observed in placental mononuclear cells.

Human fetal tissues examined (E12-E16 weeks) include: Placenta, umbilical cord, liver, kidney, adrenals, thyroid, lungs, heart, great vessels, oesophagus, stomach, small intestine, 15 spleen, thymus, pancreas, brain, eye, spinal cord, body wall, pelvis and lower limb.

Adult human tissues examined: Kidney (normal and end-stage), adrenal, myocardium, aorta, spleen, lymph node, gall bladder, pancreas, lung, skin, eye (inc. retina), prostate, bladder, liver (normal, cirrhotic, acute failure).

Non-human primate tissues examined:

20 Chimp Tissues: Salivary gland, stomach, thyroid, parathyroid, skin, thymus, ovary, lymph node.

Rhesus Monkey Tissues: Cerebral cortex, hippocampus, cerebellum, penis.

(13) DNA34431-1177 (PRO263)

Widepread expression in human fetal tissues and placenta over mononuclear cells, probably 25 macrophages +/- lymphocytes. The cellular distribution follows a perivascular pattern in many tissues. Strong expression also seen in epithelial cells of the fetal adrenal cortex. All adult tissues were negative.

Fetal tissues examined (E12-E16 weeks) include: Placenta, umbilical cord, liver, kidney, adrenals, thyroid, lungs, heart, great vessels, oesophagus, stomach, small intestine, spleen, thymus, pancreas, brain, eye, spinal cord, body wall, pelvis and lower limb.

30 Adult tissues examined: Liver, kidney, adrenal, spleen, lymph node, pancreas, lung, skin, cerebral cortex (rm), hippocampus(rm), cerebellum(rm), bladder, stomach, colon and colonic carcinoma. Acetominophen induced liver injury and hepatic cirrhosis.

A secondary screen evidenced expression over stromal mononuclear cells probably histiocytes.

35 (14) DNA38268-1188 (PRO295)

High expression over ganglion cells in human fetal spinal ganglia and over large neurones in the anterior horns of the developing spinal cord. In the adult there is expression in the chimp adrenal medulla (neural), neurones

of the rhesus monkey brain (hippocampus [+ + +] and cerebral cortex) and neurones in ganglia in the normal adult human prostate (the only section that contains ganglion cells, ie expression in this cell type is presumed NOT to be confined to the prostate). All other tissues negative.

Human fetal tissues examined (E12-E16 weeks) include: Placenta, umbilical cord, liver, kidney, adrenals, thyroid, lungs, great vessels, stomach, small intestine, spleen, thymus, pancreas, brain, eye, spinal cord, body

5 wall, pelvis, testis and lower limb.

Adult human tissues examined: Kidney (normal and end-stage), adrenal, spleen, lymph node, pancreas, lung, eye (inc. retina), bladder, liver (normal, cirrhotic, acute failure).

Non-human primate tissues examined:

Chimp Tissues: adrenal

10 Rhesus Monkey Tissues: Cerebral cortex, hippocampus, cerebellum.

Table 10

40

45

50

55

Table 10 (cont')

```

/*
*/
#include <stdio.h>
#include <ctype.h>
5
#define MAXJMP 16 /* max jumps in a diag */
#define MAXGAP 24 /* don't continue to penalize gaps larger than this */
#define JMPSS 1024 /* max jmps in an path */
#define MX 4 /* save if there's at least MX-1 bases since last jmp */
10
#define DMAT 3 /* value of matching bases */
#define DMIS 0 /* penalty for mismatched bases */
#define DINS0 8 /* penalty for a gap */
#define DINS1 1 /* penalty per base */
15
#define PINS0 8 /* penalty for a gap */
#define PINS1 4 /* penalty per residue */

20
struct jmp {
    short n[MAXJMP]; /* size of jmp (neg for delay) */
    unsigned short x[MAXJMP]; /* base no. of jmp in seq x */
    /* limits seq to 2^16 - 1 */
};

25
struct diag {
    int score; /* score at last jmp */
    long offset; /* offset of prev block */
    short ijmp; /* current jmp index */
    struct jmp jp; /* list of jmps */
};

30
struct path {
    int spc; /* number of leading spaces */
    short n[JMPSS]; /* size of jmp (gap) */
    int x[JMPSS]; /* loc of jmp (last elem before gap) */
};

35
char ofile; /* output file name */
char *namex[2]; /* seq names: getseqs() */
char *prog; /* prog name for err msgs */
char *seqx[2]; /* seqs: getseqs() */
40
int dmax; /* best diag: nw() */
int dmax0; /* final diag */
int dna; /* set if dna: main() */
int endgaps; /* set if penalizing end gaps */
int gapx, gapy; /* total gaps in seqs */
45
int len0, len1; /* seq lens */
int ngapx, ngapy; /* total size of gaps */
int smax; /* max score: nw() */
int *xbm; /* bitmap for matching */
long offset; /* current offset in jmp file */
50
struct diag *dx; /* holds diagonals */
struct path pp[2]; /* holds path for seqs */

55
char *calloc(), *malloc(), *index(), *strcpy();
char *getseq(), *g_malloc();

```

Table 10 (cont')

```

/* Needleman-Wunsch alignment program
 *
 * usage: progs file1 file2
 * where file1 and file2 are two dna or two protein sequences.
5  * The sequences can be in upper- or lower-case and may contain ambiguity
 * Any lines beginning with ';' or '<' are ignored
 * Max file length is 65535 (limited by unsigned short x in the jmp struct)
 * A sequence with 1/3 or more of its elements ACGTU is assumed to be DNA
 * Output is in the file "align.out"
10 *
 * The program may create a tmp file in /tmp to hold info about traceback.
 * Original version developed under BSD 4.3 on a vax 8650
 */
#include "nw.h"
15 #include "day.h"

static _dbval[26] = {
    1,14,2,13,0,0,4,11,0,0,12,0,3,15,0,0,0,5,6,8,8,7,9,0,10,0
};
20 static _pbval[26] = {
    1, 2|(1<<('D'-'A'))|(1<<('N'-'A')), 4, 8, 16, 32, 64,
    128, 256, 0xFFFFFFFF, 1<<10, 1<<11, 1<<12, 1<<13, 1<<14,
    1<<15, 1<<16, 1<<17, 1<<18, 1<<19, 1<<20, 1<<21, 1<<22,
    1<<23, 1<<24, 1<<25|(1<<('E'-'A'))|(1<<('Q'-'A'))
};
25 main(ac, av)
30     int      ac;
     char    *av[];
{
    prog = av[0];
    if (ac != 3) {
        35     fprintf(stderr, "usage: %s file1 file2\n", prog);
        fprintf(stderr, "where file1 and file2 are two dna or two protein sequences.\n");
        fprintf(stderr, "The sequences can be in upper- or lower-case\n");
        fprintf(stderr, "Any lines beginning with ';' or '<' are ignored\n");
        fprintf(stderr, "Output is in the file \"align.out\"\n");
        exit(1);
    }
    40     namex[0] = av[1];
    namex[1] = av[2];
    seqx[0] = getseq(namex[0], &len0);
    seqx[1] = getseq(namex[1], &len1);
    xbm = (dma)? _dbval : _pbval;
    45
    endgaps = 0;           /* 1 to penalize endgaps */
    ofile = "align.out";   /* output file */
    50     nw();           /* fill in the matrix, get the possible jmps */
    readjmps();            /* get the actual jmps */
    print();               /* print stats, alignment */
    cleanup(0);            /* unlink any tmp files */
55 }

```

Table 10 (cont')

```

/* do the alignment, return best score: main()
 * dna: values in Fitch and Smith, PNAS, 80, 1382-1386, 1983
 * pro: PAM 250 values
 5  * When scores are equal, we prefer mismatches to any gap, prefer
 * a new gap to extending an ongoing gap, and prefer a gap in seqx
 * to a gap in seq y.
 */

```

```

nw()
{
10  char      *px, *py;          /* seqs and ptrs */
    int       *ndely, *dely;    /* keep track of dely */
    int       ndelx, delx;    /* keep track of delx */
    int       *tmp;           /* for swapping row0, row1 */
    int       mis;            /* score for each type */
15  int       ins0, ins1;    /* insertion penalties */
    register id;            /* diagonal index */
    register ij;            /* jmp index */
    register *col0, *coll;   /* score for curr, last row */
    register xx, yy;        /* index into seqs */
20
    dx = (struct diag *)g_calloc("to get diags", len0+len1+1, sizeof(struct diag));

    ndely = (int *)g_calloc("to get ndely", len1+1, sizeof(int));
    dely = (int *)g_calloc("to get dely", len1+1, sizeof(int));
25  col0 = (int *)g_calloc("to get col0", len1+1, sizeof(int));
    col1 = (int *)g_calloc("to get col1", len1+1, sizeof(int));
    ins0 = (dna)? DINS0 : PINS0;
    ins1 = (dna)? DINS1 : PINS1;

30  smax = -10000;
    if (endgaps) {
        for (col0[0] = dely[0] = -ins0, yy = 1; yy <= len1; yy++) {
            col0[yy] = dely[yy] = col0[yy-1] - ins1;
            ndely[yy] = yy;
        }
        col0[0] = 0;      /* Waterman Bull Math Biol 84 */
    }
35  else
        for (yy = 1; yy <= len1; yy++)
            dely[yy] = -ins0;

40
    /* fill in match matrix
     */
    for (px = seqx[0], xx = 1; xx <= len0; px++, xx++) {
        /* initialize first entry in col
         */
        if (endgaps) {
            if (xx == 1)
                col1[0] = delx = -(ins0+ins1);
50        else
            col1[0] = delx = col0[0] - ins1;
            ndelx = xx;
        }
        else {
            col1[0] = 0;
            delx = -ins0;
            ndelx = 0;
        }
    }
55
60

```

nw

Table 10 (cont?)

...nw

```

for (py = seqx[1], yy = 1; yy <= len1; py++, yy++) {
    mis = col0[yy-1];
    if (dna)
        mis += (xbm[*px-'A']&xbm[*py-'A'])? DMAT : DMIS;
    else
        mis += _day[*px-'A'][*py-'A'];

    /* update penalty for del in x seq;
     * favor new del over ongong del
     * ignore MAXGAP if weighting endgaps
     */
    if (endgaps || ndely[yy] < MAXGAP) {
        if (col0[yy] - ins0 >= dely[yy]) {
            dely[yy] = col0[yy] - (ins0+ins1);
            ndely[yy] = 1;
        } else {
            dely[yy] -= ins1;
            ndely[yy]++;
        }
    } else {
        if (col0[yy] - (ins0+ins1) >= dely[yy]) {
            dely[yy] = col0[yy] - (ins0+ins1);
            ndely[yy] = 1;
        } else
            ndely[yy]++;
    }

    /* update penalty for del in y seq;
     * favor new del over ongong del
     */
    if (endgaps || ndelx < MAXGAP) {
        if (col1[yy-1] - ins0 >= delx) {
            delx = col1[yy-1] - (ins0+ins1);
            ndelx = 1;
        } else {
            delx -= ins1;
            ndelx++;
        }
    } else {
        if (col1[yy-1] - (ins0+ins1) >= delx) {
            delx = col1[yy-1] - (ins0+ins1);
            ndelx = 1;
        } else
            ndelx++;
    }

    /* pick the maximum score; we're favoring
     * mis over any del and delx over dely
     */
}

```

55

60

Table 10 (cont')

...nw

```

id = xx - yy + len1 - 1;
if (mis >= delx && mis >= dely[yy])
    col1[yy] = mis;
5
else if (delx >= dely[yy]) {
    col1[yy] = delx;
    ij = dx[id].ijmp;
    if (dx[id].jp.n[0] && (!dma || (ndelx >= MAXJMP
10
&& xx > dx[id].jp.x[ij]+MX) || mis > dx[id].score+DINS0)) {
        dx[id].ijmp++;
        if (++ij >= MAXJMP) {
            writejmps(id);
            ij = dx[id].ijmp = 0;
            dx[id].offset = offset;
            offset += sizeof(struct jmp) + sizeof(offset);
15
        }
    }
    dx[id].jp.n[ij] = ndelx;
    dx[id].jp.x[ij] = xx;
    dx[id].score = delx;
20
}
else {
    col1[yy] = dely[yy];
    ij = dx[id].ijmp;
25
    if (dx[id].jp.n[0] && (!dma || (ndely[yy] >= MAXJMP
&& xx > dx[id].jp.x[ij]+MX) || mis > dx[id].score+DINS0)) {
        dx[id].ijmp++;
        if (++ij >= MAXJMP) {
            writejmps(id);
30
            ij = dx[id].ijmp = 0;
            dx[id].offset = offset;
            offset += sizeof(struct jmp) + sizeof(offset);
        }
    }
    dx[id].jp.n[ij] = -ndely[yy];
35
    dx[id].jp.x[ij] = xx;
    dx[id].score = dely[yy];
}
if (xx == len0 && yy < len1) {
40
    /* last col
    */
    if (endgaps)
        col1[yy] -= ins0+ins1*(len1-yy);
    if (col1[yy] > smax) {
        smax = col1[yy];
        dmax = id;
45
    }
}
50
if (endgaps && xx < len0)
    col1[yy-1] -= ins0+ins1*(len0-xx);
if (col1[yy-1] > smax) {
    smax = col1[yy-1];
    dmax = id;
55
}
tmp = col0; col0 = col1; col1 = tmp;
}
(void) free((char *)ndely);
60
(void) free((char *)dely);
(void) free((char *)col0);
(void) free((char *)col1);
}

```

Table 10 (cont')

```

/*
 *
 * print() -- only routine visible outside this module
 *
5  * static:
 * getmat() -- trace back best path, count matches: print()
 * pr_align() -- print alignment of described in array p[]: print()
 * dumpblock() -- dump a block of lines with numbers, stars: pr_align()
 * nums() -- put out a number line: dumpblock()
10 * putline() -- put out a line (name, [num], seq, [num]): dumpblock()
 * stars() - -put a line of stars: dumpblock()
 * stripname() -- strip any path and prefix from a seqname
 */

15 #include "nw.h"

#define SPC      3
#define P_LINE   256      /* maximum output line */
#define P_SPC    3      /* space between name or num and seq */

20 extern  _day[26][26];
int     olen;           /* set output line length */
FILE   *fx;             /* output file */

25 print()                                print
{
    int     lx, ly, firstgap, lastgap; /* overlap */

    if ((fx = fopen(ofile, "w")) == 0) {
30        fprintf(stderr, "%s: can't write %s\n", prog, ofile);
        cleanup(1);
    }
    sprintf(fx, "< first sequence: %s (length = %d)\n", namex[0], len0);
    sprintf(fx, "< second sequence: %s (length = %d)\n", namex[1], len1);
35    olen = 60;
    lx = len0;
    ly = len1;
    firstgap = lastgap = 0;
    if (dmax < len1 - 1) { /* leading gap in x */
40        pp[0].spc = firstgap = len1 - dmax - 1;
        ly -= pp[0].spc;
    }
    else if (dmax > len1 - 1) { /* leading gap in y */
45        pp[1].spc = firstgap = dmax - (len1 - 1);
        lx -= pp[1].spc;
    }
    if (dmax0 < len0 - 1) { /* trailing gap in x */
50        lastgap = len0 - dmax0 - 1;
        lx -= lastgap;
    }
    else if (dmax0 > len0 - 1) { /* trailing gap in y */
55        lastgap = dmax0 - (len0 - 1);
        ly -= lastgap;
    }
    getmat(lx, ly, firstgap, lastgap);
    pr_align();
}

```

Table 10 (cont')

```

/*
 * trace back the best path, count matches
 */
static
5  getmat(lx, ly, firstgap, lastgap)                                getmat
    int      lx, ly;          /* "core" (minus endgaps) */
    int      firstgap, lastgap; /* leading/trailing overlap */
{
    int      nm, i0, i1, siz0, siz1;
10   char     outx[32];
    double    pct;
    register   n0, n1;
    register char  *p0, *p1;

15   /* get total matches, score
 */
    i0 = i1 = siz0 = siz1 = 0;
    p0 = seqx[0] + pp[1].spc;
    p1 = seqx[1] + pp[0].spc;
20   n0 = pp[1].spc + 1;
    n1 = pp[0].spc + 1;

    nm = 0;
25   while ( *p0 && *p1 ) {
        if (siz0) {
            p1++;
            n1++;
            siz0--;
        }
30   else if (siz1) {
            p0++;
            n0++;
            siz1--;
        }
35   else {
            if (xbm[*p0-'A']&xbm[*p1-'A'])
                nm++;
            if (n0++ == pp[0].x[i0])
                siz0 = pp[0].n[i0++];
40   if (n1++ == pp[1].x[i1])
                siz1 = pp[1].n[i1++];
            p0++;
            p1++;
        }
45   }

/* pct homology:
 * if penalizing endgaps, base is the shorter seq
 * else, knock off overhangs and take shorter core
 */
50   if (endgaps)
        lx = (len0 < len1)? len0 : len1;
    else
        lx = (lx < ly)? lx : ly;
55   pct = 100.* (double)nm / (double)lx;
    sprintf(fx, "\n");
    sprintf(fx, "< %d match%s in an overlap of %d: %.2f percent similarity\n",
            nm, (nm == 1)? "" : "es", lx, pct);

```

60

Table 10 (cont')

```

    sprintf(fx, "< gaps in first sequence: %d", gapx); ...getmat
    if (gapx) {
        (void) sprintf(outx, " (%d %s%s)",
5           ngapx, (dna)? "base": "residue", (ngapx == 1)? ":" : "s");
        sprintf(fx, "%s", outx);

    sprintf(fx, ", gaps in second sequence: %d", gapy);
    if (gapy) {
        (void) sprintf(outx, " (%d %s%s",
10          ngapy, (dna)? "base": "residue", (ngapy == 1)? ":" : "s");
        sprintf(fx, "%s", outx);
    }
    if (dna)
15      sprintf(fx,
          "\n< score: %d (match = %d, mismatch = %d, gap penalty = %d + %d per base)\n",
          smax, DMAT, DMIS, DINSO, DINSI);
    else
20      sprintf(fx,
          "\n< score: %d (Dayhoff PAM 250 matrix, gap penalty = %d + %d per residue)\n",
          smax, PINS0, PINS1);
    if (endgaps)
        sprintf(fx,
25          "< endgaps penalized. left endgap: %d %s%s, right endgap: %d %s%s\n",
          firstgap, (dna)? "base" : "residue", (firstgap == 1)? ":" : "s",
          lastgap, (dna)? "base" : "residue", (lastgap == 1)? ":" : "s");
    else
        sprintf(fx, "< endgaps not penalized\n");
30
}

static nm; /* matches in core -- for checking */
static lmax; /* lengths of stripped file names */
static ij[2]; /* jmp index for a path */
static nc[2]; /* number at start of current line */
35 static ni[2]; /* current elem number -- for gapping */
static siz[2];
static char *ps[2]; /* ptr to current element */
static char *po[2]; /* ptr to next output char slot */
static char out[2][P_LINE]; /* output line */
40 static char star[P_LINE]; /* set by stars() */

/*
 * print alignment of described in struct path pp[]
 */
45 static
pr_align()
{
    int nn; /* char count */
    int more;
    register i;
50
    for (i = 0, lmax = 0; i < 2; i++) {
        nn = stripname(namex[i]);
        if (nn > lmax)
            lmax = nn;
55
        nc[i] = 1;
        ni[i] = 1;
        siz[i] = ij[i] = 0;
        ps[i] = seqx[i];
        po[i] = out[i];
60
    }
}

```

Table 10 (cont')

```

for (nn = nm = 0, more = 1; more; ) {
    for (i = more = 0; i < 2; i++) {
        /*
         * do we have more of this sequence?
         */
        5      if (!*ps[i])
                continue;

        more++;

        if (pp[i].spc) { /* leading space */
            *po[i]++ = ' ';
            pp[i].spc--;
        }
        15     else if (siz[i]) { /* in a gap */
            *po[i]++ = '-';
            siz[i]--;
        }
        20     else { /* we're putting a seq element
                    */
            *po[i] = *ps[i];
            if (islower(*ps[i]))
                *ps[i] = toupper(*ps[i]);
            po[i]++;
            ps[i]++;
            /*
             * are we at next gap for this seq?
             */
            30     if (ni[i] == pp[i].x[ij[i]]) {
                /*
                 * we need to merge all gaps
                 * at this location
                 */
                35     siz[i] = pp[i].n[ij[i]++];
                while (ni[i] == pp[i].x[ij[i]]) {
                    siz[i] += pp[i].n[ij[i]++];
                }
                ni[i]++;
            }
            40     }
        }
        45     if (++nn == olen || !more && nn) {
            dumpblock();
            for (i = 0; i < 2; i++)
                po[i] = out[i];
            nn = 0;
        }
        50     }
    }

    /*
     * dump a block of lines, including numbers, stars: pr_align()
     */
    55     static
    dumpblock()
    {
        register i;
        60     for (i = 0; i < 2; i++)
            *po[i]-- = '\0';
    }
}

```

dumpblock

Table 10 (cont')

...dumpblock

```

(void) putc('\n', fx);
for (i = 0; i < 2; i++) {
    if (*out[i] && (*out[i] != ' ' || *(po[i]) != ' ')) {
        if (i == 0)
            nums(i);
        if (i == 0 && *out[1])
            stars();
        putline(i);
        if (i == 0 && *out[1])
            fprintf(fx, star);
        if (i == 1)
            nums(i);
    }
}

/*
 * put out a number line: dumpblock()
 */
static
nums(ix)
int ix; /* index in out[] holding seq line */
25 {
    char nline[P_LINE];
    register i, j;
    register char *pn, *px, *py;

    for (pn = nline, i = 0; i < lmax+P_SPC; i++, pn++)
        *pn = ' ';
    for (i = nc[ix], py = out[ix]; *py; py++, pn++) {
        if (*py == ' ' || *py == '-')
            *pn = ' ';
        else {
            if (i%10 == 0 || (i == 1 && nc[ix] != 1)) {
                j = (i < 0)? -i : i;
                for (px = pn; j /= 10, px--)
                    *px = j%10 + '0';
                if (i < 0)
                    *px = '-';
            }
            else
                *pn = ' ';
        }
        i++;
    }
    *pn = '\0';
    nc[ix] = i;
    for (pn = nline; *pn; pn++)
        (void) putc(*pn, fx);
    (void) putc('\n', fx);
}

/*
 * put out a line (name, [num], seq, [num]): dumpblock()
 */
static
putline(ix)
int ix;
60 {

```

putline

Table 10 (cont')

```

...putline
5
int      i;
register char *px;
for (px = namex[ix], i = 0; *px && *px != ':'; px++, i++)
    (void) putc(*px, fx);
for (; i < lmax+P_SPC; i++)
    (void) putc(' ', fx);
10
/* these count from 1:
 * ni[] is current element (from 1)
 * nc[] is number at start of current line
 */
15
for (px = out[ix]; *px; px++)
    (void) putc(*px&0x7F, fx);
    (void) putc('\n', fx);
}
20
/*
 * put a line of stars (seqs always in out[0], out[1]): dumpblock()
 */
static
25 stars()
{
    int      i;
    register char *p0, *p1, cx, *px;
    if (!*out[0] || (*out[0] == ' ' && *(po[0]) == ' ') ||
        !*out[1] || (*out[1] == ' ' && *(po[1]) == ' '))
        return;
    px = star;
    for (i = lmax+P_SPC; i; i--)
        *px++ = ' ';
35
    for (p0 = out[0], p1 = out[1]; *p0 && *p1; p0++, p1++) {
        if (isalpha(*p0) && isalpha(*p1)) {
            if (xbm[*p0-'A']&xbm[*p1-'A']) {
                cx = '*';
                nm++;
            }
            else if (!dn && _day[*p0-'A'][*p1-'A'] > 0)
                cx = '.';
            else
                cx = ' ';
        }
        else
            cx = ' ';
50
        *px++ = cx;
    }
    *px++ = '\n';
    *px = '\0';
55 }

```

Table 10 (cont')

<pre> /* * strip path or prefix from pn, return len: pr_align() */ static 5 stripname(pn) char *pn; /* file name (may be path) */ { register char *px, *py; 10 py = 0; for (px = pn; *px; px++) if (*px == '/') py = px + 1; 15 if (py) (void) strcpy(pn, py); return(strlen(pn)); } 20 </pre>	stripname
<pre> 25 </pre>	
<pre> 30 </pre>	
<pre> 35 </pre>	
<pre> 40 </pre>	
<pre> 45 </pre>	
<pre> 50 </pre>	
<pre> 55 </pre>	
<pre> 60 </pre>	

Table 10 (cont')

```

/*
 * cleanup() -- cleanup any tmp file
 * getseq() -- read in seq, set dna, len, maxlen
 * g_calloc() -- calloc() with error checkin
 5  * readjmps() -- get the good jmps, from tmp file if necessary
 * writejmps() -- write a filled array of jmps to a tmp file: nw()
 */
#include "nw.h"
#include <sys/file.h>

10  char  *jname = "/tmp/homgXXXXXX";           /* tmp file for jmps */
FILE  *fj;

15  int   cleanup();                         /* cleanup tmp file */
long  lseek();

/*
 * remove any tmp file if we blow
 */
20  cleanup(i)
    int   i;
{
    if (fj)
        (void) unlink(jname);
    exit(i);
}

25  /*
 * read, return ptr to seq, set dna, len, maxlen
 * skip lines starting with ';', '<', or '>'
 * seq in upper or lower case
 */
30  char  *
getseq(file, len)
35  char  *file;   /* file name */
int   *len;     /* seq len */
{
    char  line[1024], *pseq;
    register char  *px, *py;
    int   natgc, tlen;
    FILE  *fp;

    if ((fp = fopen(file, "r")) == 0) {
        fprintf(stderr, "%s: can't read %s\n", prog, file);
45    exit(1);
    }
    tlen = natgc = 0;
    while (fgets(line, 1024, fp)) {
        if (*line == ';' || *line == '<' || *line == '>')
            continue;
        for (px = line; *px != '\n'; px++)
            if (isupper(*px) || islower(*px))
                tlen++;
    }
50    if ((pseq = malloc((unsigned)(tlen+6))) == 0) {
        fprintf(stderr, "%s: malloc() failed to get %d bytes for %s\n", prog, tlen+6, file);
        exit(1);
    }
    pseq[0] = pseq[1] = pseq[2] = pseq[3] = '\0';
55
60

```

Table 10 (cont')

```

...getseq

5      py = pseq + 4;
*len = tlen;
rewind(fp);

10     while (fgets(line, 1024, fp)) {
        if (*line == ';' || *line == '<' || *line == '>')
            continue;
        for (px = line; *px != '\n'; px++) {
            if (isupper(*px))
                *py++ = *px;
            else if (islower(*px))
                *py++ = toupper(*px);
            if (index("ATGCU", *(py-1)))
                natgc++;
        }
        *py++ = '\0';
        *py = '\0';
20     (void) fclose(fp);
        dna = natgc > (tlen/3);
        return(pseq+4);
    }

25     char * g_calloc(msg, nx, sz)
    {
        char *msg; /* program, calling routine */
        int nx, sz; /* number and size of elements */
        {
30         char *px, *calloc();
            if ((px = calloc((unsigned)nx, (unsigned)sz)) == 0) {
                if (*msg) {
                    fprintf(stderr, "%s: g_calloc() failed %s (n=%d, sz=%d)\n", prog, msg, nx, sz);
                    exit(1);
                }
            }
            return(px);
        }
    }

40     /*
41      * get final jmps from dx[] or tmp file, set pp[], reset dmax: main()
42      */
43     readjmps() readjmps
44     {
45         int fd = -1;
        int siz, i0, i1;
        register i, j, xx;

50         if (fj) {
            (void) fclose(fj);
            if ((fd = open(jname, O_RDONLY, 0)) < 0) {
                fprintf(stderr, "%s: can't open() %s\n", prog, jname);
                cleanup();
            }
        }
        for (i = i0 = i1 = 0, dmax0 = dmax, xx = len0; ; i++) {
            while (1) {
                for (j = dx[dmax].jmp; j >= 0 && dx[dmax].jp.x[j] >= xx; j--)

```

Table 10 (cont')

...readjmps

```

    if (j < 0 && dx[dmax].offset && fj) {
        (void) lseek(fd, dx[dmax].offset, 0);
        (void) read(fd, (char *)&dx[dmax].jp, sizeof(struct jmp));
        (void) read(fd, (char *)&dx[dmax].offset, sizeof(dx[dmax].offset));
        dx[dmax].ijmp = MAXJMP-1;
    }
    else
        break;
}
if (i >= JMPS) {
    fprintf(stderr, "%s: too many gaps in alignment\n", prog);
    cleanup(1);
}
if (j >= 0) {
    siz = dx[dmax].jp.n[j];
    xx = dx[dmax].jp.x[j];
    dmax += siz;
    if (siz < 0) { /* gap in second seq */
        pp[1].n[i1] = -siz;
        xx += siz;
        /* id = xx - yy + len1 - 1
        */
        pp[1].x[i1] = xx - dmax + len1 - 1;
        gapy++;
        ngapy -= siz;
    /* ignore MAXGAP when doing endgaps */
        siz = (-siz < MAXGAP || endgaps)? -siz : MAXGAP;
        i1++;
    }
    else if (siz > 0) { /* gap in first seq */
        pp[0].n[i0] = siz;
        pp[0].x[i0] = xx;
        gapx++;
        ngapx += siz;
    /* ignore MAXGAP when doing endgaps */
        siz = (siz < MAXGAP || endgaps)? siz : MAXGAP;
        i0++;
    }
}
else
    break;
}
/* reverse the order of jmps
 */
for (j = 0, i0--; j < i0; j++, i0--) {
    i = pp[0].n[j]; pp[0].n[j] = pp[0].n[i0]; pp[0].n[i0] = i;
    i = pp[0].x[j]; pp[0].x[j] = pp[0].x[i0]; pp[0].x[i0] = i;
}
for (j = 0, i1--; j < i1; j++, i1--) {
    i = pp[1].n[j]; pp[1].n[j] = pp[1].n[i1]; pp[1].n[i1] = i;
    i = pp[1].x[j]; pp[1].x[j] = pp[1].x[i1]; pp[1].x[i1] = i;
}
if (fd >= 0)
    (void) close(fd);
if (fj) {
    (void) unlink(jname);
    fj = 0;
    offset = 0;
}
}

```

Table 10 (cont?)

```

/*
 * write a filled jmp struct offset of the prev one (if any): nw()
 */
5  writejmps(ix) {
    int      ix;
{
    char    *mktemp();
10   if (!fj) {
        if (mktemp(jname) < 0) {
            fprintf(stderr, "%s: can't mktemp() %s\n", prog, jname);
            cleanup(1);
        }
15   if ((fj = fopen(jname, "w")) == 0) {
            fprintf(stderr, "%s: can't write %s\n", prog, jname);
            exit(1);
        }
20   (void) fwrite((char *)&dx[ix].jp, sizeof(struct jmp), 1, fj);
        (void) fwrite((char *)&dx[ix].offset, sizeof(dx[ix].offset), 1, fj);
}
25

30

35

40

45

50

55

60

```

Table 11

PRO	XXXXXXXXXXXXXXXXXX	(Length = 15 amino acids)
Comparison Protein	XXXXXYYYYYYY	(Length = 12 amino acids)

5 % amino acid sequence identity =

(the number of identically matching amino acid residues between the two polypeptide sequences as determined by ALIGN-2) divided by (the total number of amino acid residues of the PRO polypeptide) =

10 5 divided by 15 = 33.3%

Table 12

PRO	XXXXXXXXXX	(Length = 10 amino acids)
Comparison Protein	XXXXXXXXYYYYYYZZYZ	(Length = 15 amino acids)

5 % amino acid sequence identity =

(the number of identically matching amino acid residues between the two polypeptide sequences as determined by ALIGN-2) divided by (the total number of amino acid residues of the PRO polypeptide) =

10 5 divided by 10 = 50%

Table 13

PRO-DNA	NNNNNNNNNNNNNN	(Length = 14 nucleotides)
Comparison DNA	NNNNNNNLLLLLLL	(Length = 16 nucleotides)

5 % nucleic acid sequence identity =

(the number of identically matching nucleotides between the two nucleic acid sequences as determined by ALIGN-2) divided by (the total number of nucleotides of the PRO-DNA nucleic acid sequence) =

10 6 divided by 14 = 42.9%

Table 14

PRO-DNA	NNNNNNNNNNNN	(Length = 12 nucleotides)
Comparison DNA	NNNNLLLVV	(Length = 9 nucleotides)

5 % nucleic acid sequence identity =

(the number of identically matching nucleotides between the two nucleic acid sequences as determined by ALIGN-2) divided by (the total number of nucleotides of the PRO-DNA nucleic acid sequence) =

10 4 divided by 12 = 33.3%

Deposit of Material

The following materials have been deposited with the American Type Culture Collection, 12301 Parklawn Drive, Rockville, MD, USA (ATCC):

	<u>Material</u>	<u>ATCC Dep. No.</u>	<u>Deposit Date</u>
5	DNA32292-1131	ATCC 209258	September 16, 1997
	DNA33094-1131	ATCC 209256	September 16, 1997
	DNA33223-1136	ATCC 209264	September 16, 1997
	DNA34435-1140	ATCC 209250	September 16, 1997
	DNA27864-1155	ATCC 209375	October 16, 1997
	DNA36350-1158	ATCC 209378	October 16, 1997
10	DNA32290-1164	ATCC 209384	October 16, 1997
	DNA35639-1172	ATCC 209396	October 17, 1997
	DNA33092-1202	ATCC 209420	October 28, 1997
	DNA49435-1219	ATCC 209480	November 21, 1997
	DNA35638-1141	ATCC 209265	September 16, 1997
15	DNA32298-1132	ATCC 209257	September 16, 1997
	DNA33089-1132	ATCC 209262	September 16, 1997
	DNA33786-1132	ATCC 209253	September 16, 1997
	DNA35918-1174	ATCC 209402	October 17, 1997
	DNA37150-1178	ATCC 209401	October 17, 1997
20	DNA38260-1180	ATCC 209397	October 17, 1997
	DNA39969-1185	ATCC 209400	October 17, 1997
	DNA32286-1191	ATCC 209385	October 16, 1997
	DNA33461-1199	ATCC 209367	October 15, 1997
	DNA40628-1216	ATCC 209432	November 7, 1997
25	DNA33221-1133	ATCC 209263	September 16, 1997
	DNA33107-1135	ATCC 209251	September 16, 1997
	DNA35557-1137	ATCC 209255	September 16, 1997
	DNA34434-1139	ATCC 209252	September 16, 1997
	DNA33100-1159	ATCC 209373	October 16, 1997
30	DNA35600-1162	ATCC 209370	October 16, 1997
	DNA34436-1238	ATCC 209523	December 10, 1997
	DNA33206-1165	ATCC 209372	October 16, 1997
	DNA35558-1167	ATCC 209374	October 16, 1997
	DNA35599-1168	ATCC 209373	October 16, 1997
35	DNA36992-1168	ATCC 209382	October 16, 1997
	DNA34407-1169	ATCC 209383	October 16, 1997
	DNA35841-1173	ATCC 209403	October 17, 1997
	DNA33470-1175	ATCC 209398	October 17, 1997
	DNA34431-1177	ATCC 209399	October 17, 1997
40	DNA39510-1181	ATCC 209392	October 17, 1997
	DNA39423-1182	ATCC 209387	October 17, 1997
	DNA40620-1183	ATCC 209388	October 17, 1997
	DNA40604-1187	ATCC 209394	October 17, 1997
	DNA38268-1188	ATCC 209421	October 28, 1997
45	DNA37151-1193	ATCC 209393	October 17, 1997
	DNA35673-1201	ATCC 209418	October 28, 1997
	DNA40370-1217	ATCC 209485	November 21, 1997
	DNA42551-1217	ATCC 209483	November 21, 1997
	DNA39520-1217	ATCC 209482	November 21, 1997
50	DNA41225-1217	ATCC 209491	November 21, 1997
	DNA43318-1217	ATCC 209481	November 21, 1997
	DNA40587-1231	ATCC 209438	November 7, 1997
	DNA41338-1234	ATCC 209927	June 2, 1998
	DNA40981-1234	ATCC 209439	November 7, 1997

DNA37140-1234	ATCC 209489	November 21, 1997
DNA40982-1235	ATCC 209433	November 7, 1997
DNA41379-1236	ATCC 209488	November 21, 1997
DNA44167-1243	ATCC 209434	November 7, 1997
DNA39427-1179	ATCC 209395	October 17, 1997
5 DNA40603-1232	ATCC 209486	November 21, 1997
DNA43466-1225	ATCC 209490	November 21, 1997
DNA43046-1225	ATCC 209484	November 21, 1997
DNA35668-1171	ATCC 209371	October 16, 1997

10 These deposit were made under the provisions of the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purpose of Patent Procedure and the Regulations thereunder (Budapest Treaty). This assures maintenance of a viable culture of the deposit for 30 years from the date of deposit. The deposits will be made available by ATCC under the terms of the Budapest Treaty, and subject to an agreement between Genentech, Inc. and ATCC, which assures permanent and unrestricted availability of the progeny of 15 the culture of the deposit to the public upon issuance of the pertinent U.S. patent or upon laying open to the public of any U.S. or foreign patent application, whichever comes first, and assures availability of the progeny to one determined by the U.S. Commissioner of Patents and Trademarks to be entitled thereto according to 35 USC § 122 and the Commissioner's rules pursuant thereto (including 37 CFR § 1.14 with particular reference to 886 OG 638).

20 The assignee of the present application has agreed that if a culture of the materials on deposit should die or be lost or destroyed when cultivated under suitable conditions, the materials will be promptly replaced on notification with another of the same. Availability of the deposited material is not to be construed as a license to practice the invention in contravention of the rights granted under the authority of any government in accordance with its patent laws.

25 The foregoing written specification is considered to be sufficient to enable one skilled in the art to practice the invention. The present invention is not to be limited in scope by the construct deposited, since the deposited embodiment is intended as a single illustration of certain aspects of the invention and any constructs that are functionally equivalent are within the scope of this invention. The deposit of material herein does not constitute an admission that the written description herein contained is inadequate to enable the practice of any 30 aspect of the invention, including the best mode thereof, nor is it to be construed as limiting the scope of the claims to the specific illustrations that it represents. Indeed, various modifications of the invention in addition to those shown and described herein will become apparent to those skilled in the art from the foregoing description and fall within the scope of the appended claims.

WHAT IS CLAIMED IS:

1. Isolated nucleic acid having at least 80% sequence identity to a nucleotide sequence that encodes a polypeptide comprising an amino acid sequence selected from the group consisting of the amino acid sequence shown in Figure 2 (SEQ ID NO:2), Figure 4 (SEQ ID NO:4), Figure 6 (SEQ ID NO:12), Figure 9 (SEQ ID NO:18), Figure 11 (SEQ ID NO:23), Figure 13 (SEQ ID NO:28), Figure 15 (SEQ ID NO:34), Figure 5 17 (SEQ ID NO:39), Figure 19 (SEQ ID NO:49), Figure 22 (SEQ ID NO:59), Figure 24 (SEQ ID NO:64), Figure 26 (SEQ ID NO:69), Figure 28 (SEQ ID NO:71), Figure 30 (SEQ ID NO:73), Figure 32 (SEQ ID NO:84), Figure 34 (SEQ ID NO:91), Figure 36 (SEQ ID NO:96), Figure 38 (SEQ ID NO:104), Figure 40 (SEQ ID NO:109), Figure 42 (SEQ ID NO:114), Figure 44 (SEQ ID NO:119), Figure 46 (SEQ ID NO:127), Figure 48 (SEQ ID NO:132), Figure 50 (SEQ ID NO:137), Figure 52 (SEQ ID NO:142), Figure 54 (SEQ ID NO:148), Figure 10 56 (SEQ ID NO:153), Figure 58 (SEQ ID NO:159), Figure 60 (SEQ ID NO:164), Figure 62 (SEQ ID NO:170), Figure 64 (SEQ ID NO:175), Figure 66 (SEQ ID NO:177), Figure 68 (SEQ ID NO:185), Figure 70 (SEQ ID NO:190), Figure 72 (SEQ ID NO:195), Figure 74 (SEQ ID NO:201), Figure 76 (SEQ ID NO:207), Figure 78 (SEQ ID NO:213), Figure 80 (SEQ ID NO:221), Figure 82 (SEQ ID NO:227), Figure 84 (SEQ ID NO:236), Figure 86 (SEQ ID NO:245), Figure 88 (SEQ ID NO:250), Figure 90 (SEQ ID NO:255), Figure 92 15 (SEQ ID NO:257), Figure 94 (SEQ ID NO:259), Figure 96 (SEQ ID NO:261), Figure 98 (SEQ ID NO:263), Figure 100 (SEQ ID NO:285), Figure 102 (SEQ ID NO:290), Figure 104 (SEQ ID NO:292), Figure 106 (SEQ ID NO:294), Figure 108 (SEQ ID NO:310), Figure 110 (SEQ ID NO:315), Figure 112 (SEQ ID NO:320), Figure 114 (SEQ ID NO:325), Figure 116 (SEQ ID NO:332), Figure 118 (SEQ ID NO:339), Figure 120 (SEQ ID NO:341) and Figure 122 (SEQ ID NO:377).
- 20 2. The nucleic acid of Claim 1, wherein said nucleotide sequence comprises a nucleotide sequence selected from the group consisting of the sequence shown in Figure 1 (SEQ ID NO:1), Figure 3 (SEQ ID NO:3), Figure 5 (SEQ ID NO:11), Figure 8 (SEQ ID NO:17), Figure 10 (SEQ ID NO:22), Figure 12 (SEQ ID NO:27), Figure 14 (SEQ ID NO:33), Figure 16 (SEQ ID NO:38), Figure 18 (SEQ ID NO:48), Figure 21 (SEQ ID NO:58), Figure 23 (SEQ ID NO:63), Figure 25 (SEQ ID NO:68), Figure 27 (SEQ ID NO:70), Figure 29 (SEQ ID NO:72), Figure 31 (SEQ ID NO:83), Figure 33 (SEQ ID NO:90), Figure 35 (SEQ ID NO:95), Figure 37 (SEQ ID NO:103), Figure 39 (SEQ ID NO:108), Figure 41 (SEQ ID NO:113), Figure 43 (SEQ ID NO:118), Figure 45 (SEQ ID NO:126), Figure 47 (SEQ ID NO:131), Figure 49 (SEQ ID NO:136), Figure 51 (SEQ ID NO:141), Figure 53 (SEQ ID NO:147), Figure 55 (SEQ ID NO:152), Figure 57 (SEQ ID NO:158), Figure 59 25 (SEQ ID NO:163), Figure 61 (SEQ ID NO:169), Figure 63 (SEQ ID NO:174), Figure 65 (SEQ ID NO:176), Figure 67 (SEQ ID NO:184), Figure 69 (SEQ ID NO:189), Figure 71 (SEQ ID NO:194), Figure 73 (SEQ ID NO:200), Figure 75 (SEQ ID NO:206), Figure 77 (SEQ ID NO:212), Figure 79 (SEQ ID NO:220), Figure 81 (SEQ ID NO:226), Figure 83 (SEQ ID NO:235), Figure 85 (SEQ ID NO:244), Figure 87 (SEQ ID NO:249), Figure 89 (SEQ ID NO:254), Figure 91 (SEQ ID NO:256), Figure 93 (SEQ ID NO:258), Figure 95 (SEQ ID NO:260), Figure 97 (SEQ ID NO:262), Figure 99 (SEQ ID NO:284), Figure 101 (SEQ ID NO:289), Figure 103 (SEQ ID NO:291), Figure 105 (SEQ ID NO:293), Figure 107 (SEQ ID NO:309), Figure 109 (SEQ ID NO:314), Figure 111 (SEQ ID NO:319), Figure 113 (SEQ ID NO:324), Figure 115 (SEQ ID NO:331), Figure 30 35

117 (SEQ ID NO:338), Figure 119 (SEQ ID NO:340) and Figure 121 (SEQ ID NO:376), or the complement thereof.

3. The nucleic acid of Claim 1, wherein said nucleotide sequence comprises a nucleotide sequence selected from the group consisting of the full-length coding sequence of the sequence shown in Figure 1 (SEQ ID NO:1), Figure 3 (SEQ ID NO:3), Figure 5 (SEQ ID NO:11), Figure 8 (SEQ ID NO:17), Figure 10 (SEQ ID NO:22), Figure 12 (SEQ ID NO:27), Figure 14 (SEQ ID NO:33), Figure 16 (SEQ ID NO:38), Figure 18 (SEQ ID NO:48), Figure 21 (SEQ ID NO:58), Figure 23 (SEQ ID NO:63), Figure 25 (SEQ ID NO:68), Figure 27 (SEQ ID NO:70), Figure 29 (SEQ ID NO:72), Figure 31 (SEQ ID NO:83), Figure 33 (SEQ ID NO:90), Figure 35 (SEQ ID NO:95), Figure 37 (SEQ ID NO:103), Figure 39 (SEQ ID NO:108), Figure 41 (SEQ ID NO:113), Figure 43 (SEQ ID NO:118), Figure 45 (SEQ ID NO:126), Figure 47 (SEQ ID NO:131), Figure 49 (SEQ ID NO:136), Figure 51 (SEQ ID NO:141), Figure 53 (SEQ ID NO:147), Figure 55 (SEQ ID NO:152), Figure 57 (SEQ ID NO:158), Figure 59 (SEQ ID NO:163), Figure 61 (SEQ ID NO:169), Figure 63 (SEQ ID NO:174), Figure 65 (SEQ ID NO:176), Figure 67 (SEQ ID NO:184), Figure 69 (SEQ ID NO:189), Figure 71 (SEQ ID NO:194), Figure 73 (SEQ ID NO:200), Figure 75 (SEQ ID NO:206), Figure 77 (SEQ ID NO:212), Figure 79 (SEQ ID NO:220), Figure 81 (SEQ ID NO:226), Figure 83 (SEQ ID NO:235), Figure 85 (SEQ ID NO:244), Figure 87 (SEQ ID NO:249), Figure 89 (SEQ ID NO:254), Figure 91 (SEQ ID NO:256), Figure 93 (SEQ ID NO:258), Figure 95 (SEQ ID NO:260), Figure 97 (SEQ ID NO:262), Figure 99 (SEQ ID NO:284), Figure 101 (SEQ ID NO:289), Figure 103 (SEQ ID NO:291), Figure 105 (SEQ ID NO:293), Figure 107 (SEQ ID NO:309), Figure 109 (SEQ ID NO:314), Figure 111 (SEQ ID NO:319), Figure 113 (SEQ ID NO:324), Figure 115 (SEQ ID NO:331), Figure 117 (SEQ ID NO:338), Figure 119 (SEQ ID NO:340) and Figure 121 (SEQ ID NO:376), or the complement thereof.

4. Isolated nucleic acid which comprises the full-length coding sequence of the DNA deposited under accession number ATCC 209258, ATCC 209256, ATCC 209264, ATCC 209250, ATCC 209375, ATCC 209378, ATCC 209384, ATCC 209396, ATCC 209420, ATCC 209480, ATCC 209265, ATCC 209257, ATCC 209262, ATCC 209253, ATCC 209402, ATCC 209401, ATCC 209397, ATCC 209400, ATCC 209385, ATCC 209367, ATCC 209432, ATCC 209263, ATCC 209251, ATCC 209255, ATCC 209252, ATCC 209373, ATCC 209370, ATCC 209523, ATCC 209372, ATCC 209374, ATCC 209373, ATCC 209382, ATCC 209383, ATCC 209403, ATCC 209398, ATCC 209399, ATCC 209392, ATCC 209387, ATCC 209388, ATCC 209394, ATCC 209421, ATCC 209393, ATCC 209418, ATCC 209485, ATCC 209483, ATCC 209482, ATCC 209491, ATCC 209481, ATCC 209438, ATCC 209927, ATCC 209439, ATCC 209489, ATCC 209433, ATCC 209488, ATCC 209434, ATCC 209395, ATCC 209486, ATCC 209490, ATCC 209484 or ATCC 209371.

5. A vector comprising the nucleic acid of Claim 1.

35

6. The vector of Claim 5 operably linked to control sequences recognized by a host cell transformed with the vector.

7. A host cell comprising the vector of Claim 5.
8. The host cell of Claim 7 wherein said cell is a CHO cell.
9. The host cell of Claim 7 wherein said cell is an *E. coli*.
- 5 10. The host cell of Claim 7 wherein said cell is a yeast cell.
11. A process for producing a PRO polypeptides comprising culturing the host cell of Claim 7 under conditions suitable for expression of said PRO polypeptide and recovering said PRO polypeptide from the 10 cell culture.
12. Isolated native sequence PRO polypeptide having at least 80% sequence identity to an amino acid sequence selected from the group consisting of the amino acid sequence shown in Figure 2 (SEQ ID NO:2), Figure 4 (SEQ ID NO:4), Figure 6 (SEQ ID NO:12), Figure 9 (SEQ ID NO:18), Figure 11 (SEQ ID NO:23), Figure 13 (SEQ ID NO:28), Figure 15 (SEQ ID NO:34), Figure 17 (SEQ ID NO:39), Figure 19 (SEQ ID NO:49), Figure 22 (SEQ ID NO:59), Figure 24 (SEQ ID NO:64), Figure 26 (SEQ ID NO:69), Figure 28 (SEQ ID NO:71), Figure 30 (SEQ ID NO:73), Figure 32 (SEQ ID NO:84), Figure 34 (SEQ ID NO:91), Figure 36 (SEQ ID NO:96), Figure 38 (SEQ ID NO:104), Figure 40 (SEQ ID NO:109), Figure 42 (SEQ ID NO:114), Figure 44 (SEQ ID NO:119), Figure 46 (SEQ ID NO:127), Figure 48 (SEQ ID NO:132), Figure 50 (SEQ ID NO:137), Figure 52 (SEQ ID NO:142), Figure 54 (SEQ ID NO:148), Figure 56 (SEQ ID NO:153), Figure 58 (SEQ ID NO:159), Figure 60 (SEQ ID NO:164), Figure 62 (SEQ ID NO:170), Figure 64 (SEQ ID NO:175), Figure 66 (SEQ ID NO:177), Figure 68 (SEQ ID NO:185), Figure 70 (SEQ ID NO:190), Figure 72 (SEQ ID NO:195), Figure 74 (SEQ ID NO:201), Figure 76 (SEQ ID NO:207), Figure 78 (SEQ ID NO:213), Figure 80 (SEQ ID NO:221), Figure 82 (SEQ ID NO:227), Figure 84 (SEQ ID NO:236), Figure 86 (SEQ ID NO:245), Figure 88 (SEQ ID NO:250), Figure 90 (SEQ ID NO:255), Figure 92 (SEQ ID NO:257), Figure 94 (SEQ ID NO:259), Figure 96 (SEQ ID NO:261), Figure 98 (SEQ ID NO:263), Figure 100 (SEQ ID NO:285), Figure 102 (SEQ ID NO:290), Figure 104 (SEQ ID NO:292), Figure 106 (SEQ ID NO:294), Figure 108 (SEQ ID NO:310), Figure 110 (SEQ ID NO:315), Figure 112 (SEQ ID NO:320), Figure 114 (SEQ ID NO:325), Figure 116 (SEQ ID NO:332), Figure 118 (SEQ ID NO:339), Figure 120 (SEQ ID NO:341) and Figure 122 (SEQ ID NO:377).
13. Isolated PRO polypeptide having at least 80% sequence identity to the amino acid sequence encoded by the nucleotide deposited under accession number ATCC 209258, ATCC 209256, ATCC 209264, ATCC 209250, ATCC 209375, ATCC 209378, ATCC 209384, ATCC 209396, ATCC 209420, ATCC 209480, ATCC 209265, ATCC 209257, ATCC 209262, ATCC 209253, ATCC 209402, ATCC 209401, ATCC 209397, ATCC 209400, ATCC 209385, ATCC 209367, ATCC 209432, ATCC 209263, ATCC 209251, ATCC 209255, ATCC 209252, ATCC 209373, ATCC 209370, ATCC 209523, ATCC 209372, ATCC 209374, ATCC 209373,

ATCC 209382, ATCC 209383, ATCC 209403, ATCC 209398, ATCC 209399, ATCC 209392, ATCC 209387, ATCC 209388, ATCC 209394, ATCC 209421, ATCC 209393, ATCC 209418, ATCC 209485, ATCC 209483, ATCC 209482, ATCC 209491, ATCC 209481, ATCC 209438, ATCC 209927, ATCC 209439, ATCC 209489, ATCC 209433, ATCC 209488, ATCC 209434, ATCC 209395, ATCC 209486, ATCC 209490, ATCC 209484 or ATCC 209371.

5

14. A chimeric molecule comprising a polypeptide according to Claim 12 fused to a heterologous amino acid sequence.

15. The chimeric molecule of Claim 14 wherein said heterologous amino acid sequence is an 10 epitope tag sequence.

16. The chimeric molecule of Claim 14 wherein said heterologous amino acid sequence is a Fc region of an immunoglobulin.

15 17. An antibody which specifically binds to a PRO polypeptide according to Claim 12.

18. The antibody of Claim 17 wherein said antibody is a monoclonal antibody.

FIGURE 1

ACTGCACCTCGGTTCTATCGATTGAATTCCCCGGGGATCCTCTAGAGATCCCTCGACCTCGA
CCCACCGCGTCCGGGCCGGAGCAGCACGGCCGCAGGACCTGGAGCTCCGGCTCGCTTCCCG
CAGCGCTACCCGCCATGCGCCTGCCGCCGGCGCTGGGGCTCCTGCCGCTTGCTG
CTGCTGCCGCCGCCGGAGGCCAAGAACGCGACGCCCTGCCACCGTGCCTGGGGCT
GGTGGACAAGTTAACCAAGGGATGGTGGACACCGCAAAGAACAACTTGGCGGCGGGAAACA
CGGCTTGGAGGAAAAGACGCTGTCCAAGTACGAGTCCAGCGAGATTGCCCTGCTGGAGATC
CTGGAGGGCTGTGCGAGAGCAGCAGTCGAATGCAATCAGATGCTAGAGCGCAGGAGGA
GCACCTGGAGGCCTGGTGGCTGCAGCTGAAGAGCGAATATCCTGACTTATTGAGTGGTTTT
GTGTGAAGACACTGAAAGTGTGCTGCTCTCCAGGAACCTACGGTCCCAGTCTCGCATGC
CAGGGGGATCCCAGAGGCCCTGCAGCGGAATGGCCACTGCAGCGGAGATGGAGCAGACA
GGCGACGGGTCTGCCGGTACATGGGTACCAGGGCCGCTGTGCACTGACTGCATGG
ACGGCTACTTCAGCTCGCTCCGAACGAGACCCACAGCATCTGCACAGCCTGTGACGAGTCC
TGCAAGACGTGCTGGGCCTGACCAACAGAGACTGCGCGAGTGTGAAGTGGCTGGTGCT
GGACGAGGGCGCCTGTGGATGTGGACGAGTGTGCGGCCAGCCCTCCCTGCAGCGCTG
CGCAGTTCTGTAAGAACGCCAACGGCTCCTACACGTGCGAAGAGTGTGACTCCAGCTGTG
GGCTGCACAGGGGAAGGCCAGGAAACTGTAAAGAGTGTATCTCTGGCTACGCGAGGGAGCA
CGGACAGTGTGCAAGATGTGGACGAGTGTCACTAGCAGAAAAACCTGTGAGGAAAAACG
AAAATGCTACAATACTCCAGGGAGCTACGTCTGTGTGTCCTGACGGCTTCGAAGAACG
GAAGATGCCCTGTGCGCCGGCAGAGGCTGAAGCCACAGAACGGAGAAAGCCGACACAGCT
GCCCTCCCGCGAAGACCTGTAATGTGCCGGACTTACCCCTAAATTATTCAAGAAGGATGTCC
CGTGGAAAATGTGGCCCTGAGGATGCCGTCTCTGCAGTGGACAGCGCGGGGAGAGGCTGC
CTGCTCTCTAACGGTTGATTCTCATTGTCCTTAAACAGCTGCATTCTGGTTGTTCTTA
AACAGACTTGTATTTGATACAGTTCTTGTAATAAAATTGACCATTGTAGGTAATCAGG
AGGAAAAAAGGGCGCCGCGACTCTAGAGTCGACCTGCAGAACG
TTGGCCGCCATGCCCAACTGTTATTGCAGCTTAAATGGTACAAATAAGCAATAGCA
TCACAAATTTCACAAATAAGCATTTCACTGCATTCTAGTTGTGGTTGTTGTCACAAACTC
ATCAATGTATCTTATCATGTCGGATCGGAATTAAATTGGCGCAGCACCATGGCCTGAAAT
AACCTCTGAAAGAGGAACCTGGTAGGTACCTCTGAGGCGGAAAGAACCCAGCTGTGGAATG
TGTGTCAGTTAGGGTGTGGAAAGTCCCCAGGCTCCAGCAGGCAGAAGTATGCAAGCATGC
ATCTCAATTAGTCAGCAACCCAGTTT

FIGURE 2

><subunit 1 of 1, 353 aa, 0 stop

><MW: 38192, pI: 4.53, NX(S/T): 2

MRLPRRAALGLPLLLLPPAPEAAKKPTPCRCRGLVDKFNQGMVDTAKKNFGGGNTAEEKTLSKYESSEIRL
LEILEGLCESSDFECNQMLEAQEEHLEAWWLQLKSEYPDLPFWFCVKTLKVCCSPGTYGPDCLACQGGSQRPCSG
NGHCSGDGSRQGDGSCRCHMGYQGPLCTDCMDGYFSSLRNEHTSICTACDESCKTCSGLTNRDCGECEVGWLDE
GACVDVDECAAEPPPCSAAQFCKNANGSYTCEECDSSCVGCTGEGPGNCKECISGYAREHGQCADVDECSLAEKT
CVRKNENCYNTPGSYVCVCPDGFEETEDACVPPAEEATEGESPTQLPSREDL

Signal peptide:

amino acids 1-24

N-glycosylation sites.

amino acids 190-194 and 251-255

Glycosaminoglycan attachment sites.

amino acids 149-153 and 155-159

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 26-30

Casein kinase II phosphorylation sites.

amino acids 58-62, 66-70, 86-90, 197-201, 210-214, 255-259, 295-299, 339-343

and 349-353

Tyrosine kinase phosphorylation site.

amino acids 303-310

N-myristoylation sites.

amino acids 44-50, 54-60, 55-61, 81-87, 150-156, 158-164, 164-170, 252-258 and
313-319

Aspartic acid and asparagine hydroxylation site.

amino acids 308-320

EGF-like domain cysteine pattern signature.

amino acids 166-178

Leucine zipper pattern.

amino acids 94-116

FIGURE 3

CAGGTCCAAC TG CAC CT CG GT TCT AT CG AT TG AAT TCCC GG GG AT CCT CT AG AG AT CCT C
GAC CT CG ACC CAC CG TCC CC AGG CG AGG CG AC CG CCC AG CG CT AA AC GG AA CA
GCC CT GG CT GAG GG AG CT CG AG CG CAG AG TAT CT GA CG CG CC AG GT TG CG TAG GT CG
GC AC GAGG AG GT TT CC CG AG CG AGG AG GT CT GA CG AC AT GG CC CG AGG AG CG CC TT
CCT GC CG CG CT CT GG CT CT GG AG CA T CCT CCT GT GC CT GT GC ACT GC CG CG GG AG GC
CG GG CG CG CG AGG AGG AG GC CT GT AC CT AT GG AT CG AT GT CT ACC AGG CA AG AG TA CT CA
TAGG AT TT GA AGA AG AT AT CCT GA TT GT TT CA GAG GG AAA AT GG CA CCT TT AC AC AT GAT
TT CA GAAA AG CG CA AC AG AGA AT GC CAG CT AT T CCT GT CA AT AT CC AT T CC AT GA AT TT AC
CT GG CA AG CT GC AG GG CG AG CG AGA AT ACT T CT AT GA AT T CCT GT CT TG CG CT CC CT GG GATA
AAGG CA T CAT GG CA GAT CC A ACC GT CA AT GT CCT CT GT CT GG AA CAGT GC CT CA CA AGG CA
TC AGT TG TT CA AG GT GG TT CC AT GT CT TG AAA AC AGG AT GG GT GG CG AC AT TG AAGT
GG AT GT GAT TG TT AT GA AT T CT GA AGG CA AC ACC AT T CT CCA A AC CCT CAA AT GCT AT CT
TCT TAA A AC AT GT CA AC A AG CT GA GT GC CC AGG CG GG GT CG CA A AT GG AGG CTT TG AAT
GAA AG AC GC AT CT CG AG GT GT CCT GA TGG GT CC AC GG AC CT CA CT GT GA GAA AG CC CTT
TAC CC AC GAT GT AT GA AT GG GG ACT T GT GT GACT CCT GT TT CT GC AT CT GC CC AC CT
GAT T CT AT GG AG GT GA ACT GT GA CAA AG CAA AC TG CT CA ACC AC CT GT TT AT GG AGG GAC
TG TT CT ACC CT GG AAA AT GT AT TT GC C CT CC AGG ACT AG AGG AG AG CAGT GT GAA AT CAG
CAA AT GC CC AC A ACC CT GT CG AA AT GG AG GT AA AT GC AT TG T AAA AG CAA AT GT AAGT GT
CC AA AG GT T ACC CAGG AG AC CT CT GT T CA A AG C CT GT CT CG AG C CT GG CT GT GG TG CAC AT
GG A AC CT GC CAT GA ACC CA A AA AT GC CA AT GT CA AGA AG GT GG AT GG AA AG AC AT GC
TAA A AG GT AC GA AG CC AG C CT C AT AC AT GC C CT TG AGG CC AG C AGG CG CC CAG CT CAGG CAGC
AC AC GC CT T C ACT T AAA AGG CG AGG AG CG GG CG GG AT CC AC CT GA AT CC A ATT AC AT CT GG
TGA ACT CC CG AC AT CT GA A AC GT TT AAGT TA CAC CA AG GT CAT AG C CTT GT TA AC CTT CA
TG T GT GA AT GT T CAA A AT GT T CATT AC ACT TA AGA AT ACT TG G C CT GA AT TT T ATT AG CT
TC ATT AT A A AT CA CT GAG CT GAT AT TT ACT CT C TT TT AAGT TT CT AAGT AC GT CT GT TAG
CAT GAT GG T AT AG AT TT CT GT TT CAGT GCTT GG AC AG AT TT T AT ATT GT CA ATT GA
TC AGG TT AAA AT TT CAGT GT TAG GT GG CAG AT AT TT CAAA AT TA CA AT GC AT TT AT GG
GT CT GG GG CG AG GG AA CA CT CA GAA AG GT TAA AT TT GG CAAA AT GC GT AAGT CAC A AGA AT
TT GG AT GG TG CAG T TA AT GT GA AG GT AC AG C AT TT CAG AT TT T ATT GT CAG AT AT T AG AT
GT TT GT AC AT TT T AAA AT TG CT TT AAT TT TAA ACT CT CA AT CA AT AT ATT TT GACC
TT ACC ATT AT CC AGA GAT T CAGT ATT AAA
AA A CA AT AT A A AT AT CT AA AC AC A AT GA A AT AGG A AT AT A AT GT AT GA A CT TT TG C AT
TGG C TT GA AG CA AT AT A A AT AT TT GT TAA AC A A A A C AC AG C T CT TA CCT A A A A C A T T T
A C T GT TT GT AT GT AT AAA
A
CG CC AT GG CC CA CT GT TT ATT GC AG CT T A A T G

FIGURE 4

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA33094
><subunit 1 of 1, 379 aa, 0 stop
><MW: 41528, pI: 7.97, NX(S/T): 2
MARRSAFPAAALWLWSILLCLLALRAEAGPPQEEESLYLWIDAHQARVLIGFEEDILIVSEGK
MAPFTHDFRKAQQRMPAIPVNIHSMNFTWQAAGQAEYFYEFLSLRSLDKGIMADPTVNVPPLL
GTVPHKASVVQVGFPCLGKQDGVAAFEVDVIVMNSEGNTILQTPQNAIFFKTCQQAECPGGC
RNGGFCNERRICECPDGFHGPHEKALCTPRCMNGGLCVTPGFCICPPGFYGVNCDKANCST
TCFNGGTCFYPGKICPPGLEGEQCEISKCPQPCRNGGKIGKSKCKCSKGYQGDLCSPV
EPGCGAHGTCHEPNKCQCQEGWHGRHCNKRYEASLIHALRPAGAQLRQHTPSLKKAEERRDP
PESNYIW
```

Signal peptide:

amino acids 1-28

N-glycosylation site.

amino acids 88-92, 245-249

Casein kinase II phosphorylation site.

amino acids 319-323

Tyrosine kinase phosphorylation site.

amino acids 370-378

N-myristylation sites.

amino acids 184-190, 185-191, 189-195, 315-321

ATP/GTP-binding site motif A (P-loop).

amino acids 285-293

EGF-like domain cysteine pattern signature.

amino acids 198-210, 230-242, 262-274, 294-306, 326-338

FIGURE 5

CGGACGCGTGGCGTCCGGCGTCGAGAGCCAGGAGGCGGAGGCGCGGGCCAGCCCTGGG
 CCCCAGCCCACACCTTACCAAGGGCCAGGAGCCACCATGTGGCGATGTCCACTGGGGCTAC
 TGCTGTTGCTGCCGCTGGCTGGCCACTTGGCTCTGGGTGCCAGCAGGGCTGTGGGCCGG
 GAGCTAGCACCAGGGCTGCACCTGCAGGGCATCCGGGACGCCGGAGGCGGTACTGCCAGGA
 GCAGGACCTGTGCTGCCGCGGCCGTGCCAGCAGACTGTGCCCTGCCCTACCTGGGCCATCT
 GTTACTGTGACCTCTCTGAACCGCACGGTCTCGACTGTCTGCCCTGACTTCTGGGACTTC
 TGCCTCGGCGTGCACCCCTTTCCCCGATCCAAGGATGTATGCATGGAGGTGATCTA
 TCCAGTCTGGGAAACGTACTGGGACAACGTAAACGTTGCACCTGCCAGGAGAACAGGCAGT
 GGCATGGTGGATCCAGACATGATCAAAGCCATCAACCAGGGCAACTATGGCTGGCAGGCTGG
 GAACACAGCGCCTCTGGGCATGACCTGGGATGAGGGCATTGCTACCGCCTGGCACCA
 TCCGCCATCTCCTCGGTATGAACATGCATGAAATTATACAGTGTGAACCCAGGGAG
 GTGCTTCCCACAGCCTCGAGGCCTCTGAGAAGTGGCCCAACCTGATTGATGAGCCTCTTGA
 CCAAGGCAACTGTGCAAGGCCTCTGGCCTTCTCCACAGCAGCTGTGGCATCCGATCGTGTCT
 CAATCCATTCTCTGGGACACATGACGCCCTGTCTGTCGCCAGAACCTGCTGTCTTGAC
 ACCCACCAGCAGCAGGGCTGCCGCGTGGCGTCTCGATGGTGCCTGGTGGTCTGCGTCG
 CCGAGGGGTGGTCTGACCACTGCTACCCCTCTCGGGCGTGAACGAGACGAGGCTGGCC
 CTGCGCCCCCTGTATGATGACAGCCGAGCCATGGTCGGGCAAGCGCCAGGCCACTGCC
 CACTGCCAACAGCTATGTTAATAACAATGACATCTACAGGTCACTCCTGTCTACGCC
 CGGCTCCAACGACAAGGAGATCATGAAGGAGCTGATGGAGAATGGCCCTGTCCAAGGCC
 TGGAGGTGCATGAGGACTTCTCCTATACAAGGGAGGCATCTACAGCCACACGCCAGTGAGC
 CTTGGGAGGCCAGAGAGATACCGCCGGCATGGGACCCACTCAGTCAAGATCACAGGATGGGG
 AGAGGAGACGCTGCCAGATGGAAGGACGCTAAATACTGGACTGCGGCCACTCCTGGGGCC
 CAGCCTGGGCGAGAGGGGCCACTTCCGCATCGTGCAGGGCGTCAATGAGTGCACATCGAG
 AGCTTCTGTCTGGCGTCTGGGCCCGTGGCATGGAGGACATGGTCATCACTGAGGCTG
 CGGGCACACGCCGGTCCGGCTGGATCCAGGCTAAGGGCGGCGGAAGAGGCCCAATG
 GGGCGGTGACCCAGCCTCGCCGACAGAGCCGGGCGCAGGCGGGCGCAGGGCGCTAAT
 CCCGGCGGGTCCGCTGACGCAAGGCCCGCTGGAGGCCAGGGCAGGCGAGACTGGCG
 GAGCCCCAGACCTCCCAGTGGGACGGGCAGGGCTGGCCTGGGAAGAGCACAGCTGCAG
 ATCCCAGGCCCTGGCGCCCCACTCAAGACTACCAAGCCAGGACACCTCAAGTCTCCAGC
 CCCAATACCCACCCCAATCCGTATTCTTTTTTTAGACAGGGCTTGCTCCGACTCCTGGGTTCA
 TTGCCCAGGTTGGAGTGCAGTGGCCATCAGGGCTACTGTAACCTCCGACTCCTGGGTTCA
 AGTGACCCCTCCCACCTCAGCCTCTCAAGTAGCTGGACTACAGGTGCACCAACACCTGGC
 TAATTTTGATTTTGAAAGAGGGGGTCTCACTGTGTTGCCAGGCTGGTTCAACT
 CCTGGCTCAAGCGGTCCACCTGCCCTCCCAAAGTGTGGATTGCAGGCATGAGCC
 ACTGCACCCAGCCGTATTCTTCAAGATATTATTTCTTCACTGTTAAAA
 TAAAACCAAAGTATTGATAAAAAAAA

FIGURE 6

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA33223
><subunit 1 of 1, 164 aa, 1 stop
><MW: 18359, pI: 7.45, NX(S/T): 1
MWRCPLGLLLLLPIAGHLALGAQQGRGRRELAPGLHLRGIRDAAGGRYQCQEQLCCRGRADDC
ALPYLGAICYCDLFCNRTVSDCCPDFWDFCLGVPPPFPPPIQGCMHGGRIYPVLGTYWDNCNR
CTCQENRQWHGGSRHDQSHQPGQLWLAGWEPQRILLGHDPG
```

N-glycosylation site.

amino acids 78-82, 161-165

Casein kinase II phosphorylation site.

amino acids 80-84, 117-121, 126-130, 169-173, 205-209, 296-300,
411-415

N-myristoylation site.

amino acids 21-27, 39-45, 44-50, 104-110, 160-164, 224-230,
269-275, 378-384, 442-448

Amidation site.

amino acids 26-30, 318-322

Eukaryotic thiol (cysteine) proteases histidine active site.

amino acids 398-409

FIGURE 7

AGGCTCCTGGCCCTTTCCACAGCAAGCTNTGCNATCCGATTGTTGTCCTAAATCCA
ATTCTCTGGGACACATNACGCCTGTCCTTNGCCCCAGAACCTGCTGTCTTGTACACCCAC
CAGCAGCAGGGCTGCCCGNTGGCGTCTCGATGGTGCCTGGTGGTCTCGCTGCCGAGG
GNTGGTGTCTGACCACTGCTACCCCTCTCGGCCGTGAACGAGACGAGGCTGCCCTGCGC
CCCCCTGTATGATGCACAGCCGAGCCATGGTCGGGCAAGGCCAGGCCACTGCCACTGC
CCCAACAGCTATGTTAATAACAATGACATCTACCAAGGTCACTCCTGTCTACGCCCTGGCTC
CAACGACAAGGAGATCATGAAGGAGCTGATGGAGAATGCCCTGTCCAAGCCCTCATGGAGG
TGCATGAGGACTTCTTCCTATAACAAGGGAGGCATCTACAGCCACAGCCAGTGAGCCTGGG
AGGCCAGAGAGATAACGCCGGCATGGGACCCACTCAG

FIGURE 8

GCTGCTGCCCTGTTGATGGCAGGCTTGGCCCTGCAGCCAGGCAGTGCCTGCTGTGCTACT
CCTGCAAAGCCCAGGTGAGCAACGAGGACTGCCTGCAGGTGGAGAACTGCACCCAGCTGGGG
GAGCAGTGCTGGACCGCGCGATCCGCGCAGTGGCCTCCTGACCGTCATCAGCAAAGGCTG
CAGCTTGAAC~~T~~CGTGGATGACTCACAGGACTACTACGTGGCAAGAAGAACATCACGTGCT
GTGACACCGACTTGTGCAACGCCAGCGGGGCCATGCCCTGCAGCCGGCTGCCGCCATCCTT
GCGCTGCTCCCTGCACTCGCCTGCTGCTCTGGGACCCGGCCAGCTATAGGCTCTGGGGGG
CCCCGCTGCAGCCCACACTGGGTGTGGTGCCCCAGGCCTCTGTGCCACTCCTCACAGACCTG
GCCAGTGGGAGCCTGCTGGTCTGGAGGCACATCCTAACGCAAGTCTGACCATGTATGT
CTGCACCCCTGTCCCCCACCTGACCCCTCCATGCCCTCTCAGGACTCCCACCCGGCAGA
TCAGCTCTAGTGACACAGATCCGCTGCAGATGGCCCTCCAACCCCTCTGCTGCTGTTTC
CATGGCCCAGCATTCTCCACCCCTAACCCCTGTGCTCAGGCACCTCTCCCCCAGGAAGCCTT
CCCTGCCACCCCATCTATGACTTGAGCCAGGTCTGGTCCGTGGTGTCCCCCGCACCCAGCA
GGGGACAGGCACTCAGGAGGGCCAGTAAAGGCTGAGATGAAGTGGACTGAGTAGAAACTGGA
GGACAAGAGTCACGTGAGTTCTGGAGTCTCCAGAGATGGGCCTGGAGGCCTGGAGGAA
GGGGCCAGGCCTCACATTGTGGGCTCCCTGAATGGCAGCCTGAGCACAGCGTAGGCCCTT
AATAAACACCTGTTGGATAAGCCAAAAAA

FIGURE 9

MTHRTTWARRTSRAVTPTCATPAGPMPCSRLPPSLRCSLHSACCSGDPASYRLWGAPIQPT
LGVVPQASVPLLTDLAQWEPVLVPEAHPNASLTMYVCTPVPHDPPMALSRTPTRQISSSDT
DPPADGPSNPLCCCFHGPAFSTLNPVLRHLFPQEAFPAHPIYDLSQVWSVSPAPSRGQALRRAQ

Signal peptide:

amino acids 1-47

N-glycosylation site.

amino acids 31-35, 74-78, 84-88

Casein kinase II phosphorylation site.

amino acids 22-26, 76-80

N-myristoylation site.

amino acids 56-60

Amidation site.

amino acids 70-74

FIGURE 10

CCACACGCGTCCGAACCTCTCCAGCGATGGGAGCCGCCCTGCTGCCAACCTCACTCTGT
GCTTACAGCTGCTGATTCTCTGCTGTCAAACCTCAGTACGTGAGGGACCAGGGGCCATGACC
GACCAGCTGAGCAGGCGGCAGATCCCGAGTACCAACTCTACAGCAGGACCAGTGGCAAGCA
CGTGCAGGTACCGGGCGTCGCATCTCCGCCACCGCCGAGGAACAGGAACAAGTTGCCAAGC
TCATAGTGGAGACGGACACGTTGGCAGCCGGTCGCATCAAAGGGCTGAGAGTGAGAAG
TACATCTGTATGAACAAGAGGGCAAGCTCATCGGAAGGCCAGCGGAAGAGCAAAGACTG
CGTGGTCACGGAGATCGTGCAGAGAACAACTATAACGGCCTTCCAGAACGCCGGCACGAGG
GCTGGTCATGGCCTTCACCGGGCAGGGCGGGCCAGGCTTCCGCAGCCGCCAGAAC
CAGCGCGAGGCCACTTCATCAAGGCCCTTACCAAGGCCAGCTGCCCTCCCCAACACGC
CGAGAACGAGAACAGCTGAGTTGGCTCCGCCCCACCCGCCGGACCAAGCGCACAC
GGCGGCCCTCACGTAGTCTGGGAGGCAGGGGCAGCAGCCCTGGCCGCCCTCCC
CACCCCTTCCCTCTTAATCCAAGGACTGGGCTGGGTGGCGGGAGGGGAGGCCAGATCCCC
GAGGGAGGACCTGAGGGCCGCGAAGCATCCGAGCCCCCAGCTGGGAAGGGGAGGCCGTG
CCCCAGGGCGGCTGGCACAGTCCCCCTCCGGACGGTGGCAGGCCCTGGAGAGGAACT
GAGTGTACCCCTGATCTCAGGCCACCAGCCTCTGCCGCCCTCCAGCCGGCTCTGAAGCC
CGCTGAAAGGTCAAGCGACTGAAGGCCTTGCAGACAACCGTCTGGAGGTGGCTGTCTCAAAA
TCTGCTCTCGGATCTCCCTCAGTCTGCCCTCAGCCCCAAACTCCTCTGGCTAGACTGTA
GGAAGGGACTTTGTTGTTGTTGTTGAGAAAAAGAAAGGGAGAGAGAGGAAATAG
AGGGTTGTCCACTCCTCACATTCCACGACCCAGGCCTGCACCCACCCCAACTCCCAGCCC
CGGAATAAAACCATTTCCCTGC

FIGURE 11

MGAARLLPNLTLCLQLLILCCQTQYVRDQGAMTDQLSRRQIREYQLYSRTSGKHVQVTGRRI
SATAEDGNKFAKLIVETDTFGSRVRIKGAESEKYICMNKRKGKLIKPSGKSKDCVFTEIVLE
NNYTAFQNARHEGWFMAFTRQGRPRQASRSRQNQREAHFIKRLYQGQLPFPNHAEKQQQEF
VGSAPTRRTKRTRRPQPLT

Signal peptide:

amino acids 1-22

N-glycosylation site.

amino acids 9-13, 126-130

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 60-64

Casein kinase II phosphorylation site.

amino acids 65-69

Tyrosine kinase phosphorylation site.

amino acids 39-48, 89-97

N-myristoylation site.

amino acids 69-75, 188-194

Amidation site.

amino acids 58-62

HBGF/FGF family signature.

amino acids 103-128

FIGURE 12

ACCTGCCATCACCTGTTGCCAGTGTGGAAAAATTCTCCCTGTTGAATTTTGACATGGAG
GACAGCAGCAAAGAGGGCAACACAGGCTGATAAGACCAAGAGACAGCAGGGAGATTATTTAC
CATACGCCCTCAGGACGTTCCCTCTAGCTGGAGTTCTGGACTTCACAGAACCCCATCCAGT
CATTTGATTTGCTGTTATTTTTCTTTCTTCCCACACATTGTATTATTTAT
TTCCGTACTTCAGAAATGGCCTACAGACCACAAAGTGGCCAGCCATGGGCTTTTCCT
GAAGTCTGGCTTATCATTCCCTGGGCTCTACTCACAGGTGTCACACTCCCTGGCTGCC
CTAGTGTGTGCCGCTGCACAGGAACCTTGCTACTGTAATGAGCGAAGCTGACCTCAGTG
CCTCTGGGATCCGGAGGGCGTAACCGTACTCTACCTCCACAAACAAACAAATTAAATGC
TGGATTTCCTGCAGAACTGCACAATGTACAGTCGGTGACACGGCTACCTGTATGCCAAC
AACTGGACGAATTCCCCATGAACCTCCAAAGAATGTCAGAGTTCTCCATTGCAAGGAAAAC
AAATATTCA GACCAATTTCACGGGCTGCTCTTGCCCAGCTCTGAAGCTTGAAGAGCTGACACT
GGATGACAACCTCCATATCCACAGTGGGGTGAAGACGGGCTTCCGGGAGGCTATTAGCC
TCAAATTGTTGTTTGTCTAAGAATCACCTGAGCAGTGTGCTTGGCTTGTGGAC
TTGCAAGAGCTGAGAGTGGATGAAAATCGAATTGCTGTATATCCGACATGCCCTCCAGAA
TCTCACGAGCTTGGAGCGTCTATTGTGGACGGAACCTCCTGACCAACAAAGGTATCGCCG
AGGGCACCTTCAGCCATCTCACCAAGCTCAAGGAATTTCATAATTGTACGTAATTGCTGTCC
CACCCCTCCTCCGATCTCCAGGTACGCATCTGATCAGGCTCTATTGCAAGGACAACQAGAT
AAACCACATTCCCTTGACAGCCTCTCAAATCTGCTAAGCTGGAACGGCTGGATATATCCA
ACAACCAACTGCGGATGCTGACTCAAGGGTTTGATAATCTCTCCAAACCTGAAGCAGCTC
ACTGCTCGGAATAACCCCTGGTTTGACTGCAGTATAATGGGTACAGAAATGGCTCAA
ATATATCCCTTCATCTCAACGTGGGGTTTCATGTGCCAACGGCTCTGAACAAGTCCGGG
GGATGCCGTAGGAATTAAATATGAATCTTGTCTGTCCCACACGACCCCCGGCCTG
CCTCTTCACCCCAGCCCCAAGTACAGCTCTCCGACCACCTCAGCCTCCCACCCCTCTAT
TCCAAACCCTAGCAGAAGTACACGCCTCAACTCCTACCACATCGAAACTCCACGATT
CTGACTGGATGGCAGAGAAAAGAGTGACCCCACCTATTCTGAACGGATCCAGCTCTATC
CATTTGTGAATGATACTCCATTCAAGTCAGCTGGCTCTCTCTTCAACCGTGTGGCATA
CAAACACATGGGTGAAAATGGGCACAGTTAGTAGGGGGCATGTTCAAGGAGCGCATAG
TCAGCGGTGAGAAGCAACACCTGAGCCTGGTTAACTTAGAGCCCCGATCCACCTATGGATT
TGTTTAGTGCCTGGATGCTTTAACTACCGCGCGTAGAAGACACCATTGTTCAAGG
CACCAACCCATGCCCTCTATCTGAACAAACGGCAGCAACACAGCGTCCAGCCATGAGCAGCG
CGTCCCACAGCATGGCTCCCCCTTCTGCTGGCGGGCTTGATGGGGCGCGGTGATATT
GTGCTGGTGGCTTGCTCAGCGTCTTTGCTGGCATATGCACAAAAAGGGCGCTACACCTC
CCAGAAGTGGAAATACAACCGGGGCCGGCGAAAGATGATTATGCGAGGCAGGCACCAAGA
AGGACAACCTCCATCCTGGAGATGACAGAAACAGTTTCAGATCGTCTCCTAAATAACGAT
CAACTCCTAAAGGAGATTTCAGACTGCAGCCATTACACCCAAATGGGGCATTAAATTA
CACAGACTGCCATATCCCCAACACATGCGATACTGCAACAGCAGCGTGCCAGACCTGGAGC
ACTGCCATACGTGACAGCCAGAGGCCAGCGTTATCAAGGCGACAATTAGACTCTTGAGAA
CACACTCGTGTGCACATAAGACACGCAGATTACATTGATAATGTTACACAGATGCAT
TTGTGCATTGAATACTCTGTAATTATACGGTGTACTATATAATGGGATTAAAAAAAGTG
CTATCTTCTATTCAAGTTAATTACAAACAGTTGTAACTCTTGCTTTAAATCTT

FIGURE 13

MGLQTTKWPShGAFFLKS~~W~~LI ISLGLYSQVS~~K~~LLACPSVCRC~~R~~DNFVYC~~N~~ERSLTSVPLGIP
EGVT~~V~~LYLHNNQINNAGFPAELHNQSVHTVYLYGNQLDEFPMNL~~P~~KNVRVL~~H~~QENNIQTI
SRAALAQLLKLEELHLD~~D~~NS~~S~~ISTVGVEDGAF~~R~~EAISL~~K~~LLFLSKNHLSSPV~~G~~LPV~~D~~LQELR
VDENRIAVISDMAFQNL~~T~~SLERLIVDGNLLTNKGIAEGTF~~S~~HLTKL~~K~~EFSIVRNSLSH~~PP~~D
LPGTHLIRLYLQDNQINH~~I~~PLTA~~F~~SNLRKLERLDISNNQLRMLTQGVFDNL~~S~~NLKQLTARNN
PWFCDCS~~I~~KWVTEWLKY~~I~~PSSLNV~~R~~GFMCQGPEQVRGM~~A~~VRELN~~M~~LLSCPTTPGLPLFTP
APSTASPTTQPPTLS~~I~~P~~N~~PSRSY~~T~~P~~P~~PTT~~S~~KLPT~~I~~PDWDGR~~R~~ERT~~P~~P~~I~~SERIQLSIHFVND
TSIQVSWL~~S~~LFTVMAYKLTWVKMGHSLVGGIVQERIVSGEKQHLSLVNLEPRSTYRICLVPL
DAFNYRAVEDTICSEATTHASYLNNGSNTASSHEQTTSHSMGSPFLLAGLIGGAVIFVLVVL
LSVFCWHMHKKGRYTSQWKY~~N~~RGRKDDYCEAGTKKD~~N~~SILEM~~T~~TSFQIVSLNNDQLLKG
DFRLQPIYTPNGGINYTD~~C~~HIPNNMRYCNSSVPDLEH~~C~~HT

Signal peptide:

amino acids 1-42

Transmembrane domain:

amino acids 542-561

N-glycosylation site.

amino acids 202-206, 298-302, 433-437, 521-525, 635-639, 649-653

Casein kinase II phosphorylation site.

amino acids 204-208, 407-411, 527-531, 593-597, 598-602, 651-655

Tyrosine kinase phosphorylation site.

amino acids 319-328

N-myristoylation site.

amino acids 2-8, 60-66, 149-155, 213-219, 220-226, 294-300,
522-528, 545-551, 633-639

Amidation site.

amino acids 581-585

Leucine zipper pattern.

amino acids 164-186

Phospholipase A2 aspartic acid active site.

amino acids 39-50

FIGURE 14

ACTTGGAGCAAGCGCGGGCGGAGACAGAGGCAGAGGCAGAAGCTGGGCTCCGTCCGCCTCCACGAGCG
 ATCCCCGAGGAGAGCCGCGGCCCTCGCGAGGCAGAGGCCAGAGGAAGACCCGGGTGGCTGCGCCCTGCC
 TCGCTTCCCAGGCAGGCCGGCTGCAGCCTGCCCTCTGCTCGCCTGAAAATGGAAAAGATGCTCGCAGGCT
 GCTTTCTGCTGATCCTCGGACAGATCGCCTCCCTCCGAGGGCAGGGAGCGGTACGTGGAGGTCCATCT
 CTAGGGCAGACACGCTCGGACCCACCGCAGACGCCCTCTGGAGAGTTCCTGTGAGAACAGCAGGGCAGACC
 TGGTTTCATCATTGACAGCTCGCAGTGTCAACACCCATGACTATGCAAAGGTCAAGGAGTTCATCGGGACA
 TCTTGCAATTCTTGACATTGGCCTGATGTCACCCGAGTGGGCTGCTCAATATGGCAGCACTGTCAAGAATG
 AGTTCTCCCTCAAGACCTCAAGAGGAAGTCCGAGGTGGAGCGTGTCAAGAGGATGCCATCTGCCACGG
 GCACCATGACTGGGCTGCCATCCAGTATGCCCTGAACATCGCATTCTCAGAACAGCAGAGGGGCCGGCCCTGA
 GGGAGAATGTGCCACGGGTCTATAATGATCGTACAGATGGGAGACCTCAGGACTCCGTGGCCAGGTGGCTGCTA
 AGGCACGGGACACGGGCATCCTAATCTTGGCATTGGTGTGGCCAGGTAGACTTCAACACCTGAAAGTCCATTG
 GGAGTGAGCCCCATGAGGACCATGTCCTCTGTGGCAATTTCAGCCAGATTGAGACGCTGACCTCCGTGTTCC
 AGAAGAAGTTGTGACGGCCACATGTGACGACCCCTGGAGCATAACTGTGCCACTTCTGCATCAACATCCCTG
 GCTCATACGTCTGAGGTCAAACAAGGCTACATTCTCAACTCGGATCAGACGACTTGCAAGAATCCAGGATCTGT
 GTGCCATGGAGGACCAACTGTGAGCAGCTGTGTGAATGTGCCGGCTCTCGTCTGCCAGTGCTACAGTG
 GCTACGCCCTGGCTGAGGATGGGAAGAGGTGTGGACTACTGTGCCCTCAGAAAACCACGGATGTGAAC
 ATGAGTGTGAAATGCTGATGCCCTACCTTGCCAGTGCATGAAGGATTGCTCTTAACCCAGATGAAAAAA
 CGTGCACAAGGATCAACTACTGTGCACTGAACAAACGGGCTGTGAGCATGAGTGTGCAACATGGAGGAGAGCT
 ACTACTGCGCTGCCACCGTGGCTACACTCTGGACCCCAATGGAAAACCTGAGCCAGTGGACACTGTGAC
 AGCAGGACCATGGCTGTGAGCAGCTGTGTGAACACGGAGGATTCTCGTCTGCCAGTGCTCAGAACGGCTTCC
 TCATCAACGAGGACTCAAGACCTGCTCCGGGTTGGAATTACTGCCCTGCTGAGTGAACATGGTTGTGAATACTCCT
 GTGTCAACATGGACAGATCCTTGCCCTGTCAGTGTGCTGAGGGACACGTGCTCCGAGCGATGGGAAGACGTGTG
 CAAAATTGGACTCTGTGCTTGGGGACACGGGTTGTAACATTGCTGTGAAGCAGTGAAGGATTGTTGT
 GCCAGTGCTTGAAAGTTATATACTCCGTGAAGATGGAAAAACCTGAGAACAGGAAAGATGTCGCCAAGCTATAG
 ACCATGGCTGTGAACACATTGTTGTAACAGTGCAGACTCATACACGGTGGAGTGGCTTGGAGGATTCCGGCTCG
 CTGAGGATGGGAAACGCTGCCAGAGGAAGGGATGTCTGCAAATCAACCCACCATGGCTGCAACATTTGTGTTA
 ATAATGGAATTCCACATCTGCAAATGCTCAGAGGGATTGTTGCTAGCTGAGGACCGGAAGACGGTCAAGAAAT
 GCACTGAAGGCCAATTGACCTGGTCTTGATGATGGATCCAAGAGTCTGGAGAAGAGAATTGGAGGTG
 TGAAGCAGTTGCACTGGAATTAGATTCTTGACAAATTCCCCAAAGCCCTCGAGTGGGGCTGCTCCAGT
 ATTCCACACAGGTCCACACAGAGTTCACTCTGAGAAACTTCAACTCAGCCAAAGACATGAAAAAGCCGTGGCCC
 ACATGAAATACATGGAAAGGGCTATGACTGGCTGGCCCTGAAACACATGTTGAGGAAAGTAAACCCAG
 GAGAAGGGCCAGGCCCTTCCACAAGGTGCCAGAGCAGCATTGTTCACCGACGGACGGCTCAGGATG
 ACGTCTCGAGTGGCCAGTAAAGCCAAGGCAATGGTATCACTATGTATGCTGTTGGGTAGGAAAAGCCATTG
 AGGAGGAACATACAAGAGATTGCCCTGAGGCCACAAACAAGCATCTCTTCTATGCCAGACTTCAGCACAATGG
 ATGAGATAAGTGAAGAAACTCAAGAAAGGCATCTGTGAAGCTCTAGAAGACTCCGATGGAAGACAGGACTCTCAG
 CAGGGGAACCTGCCAAACGGTCAAACAGCAACAGAACATCTGAGCAGTCACCATAAATATCCAAGACCTACTTT
 CCTGTTCTAATTGCACTGCAACACAGATATCTGTTGAAGAAGACAATCTTACGGTCTACACAAAGCTTT
 CCCATTCAACAAACCTTCAGGAAGGCCCTTGGAGAAGAAAAACACGATCAATGCAAATGTGAAAACCTTATAATGT
 TCCAGAACCTGCAAACGAAGAAGTAAGAAAATTACACAGCGTTAGAAGAAATGACACAGAGAATGGAAGGCC
 TGGAAAATGCCCTGAGATACAGATGAGATTAGAAATCGCGACACATTGTTGAGTCATTGTATCAGGATTACAAT
 GAACGCAGTGCAGAGGCCAAAGCTCAGGCTATTGTTAAATCAATAATGTTGAGTAAAACAATCAGTACTGA
 GAAACCTGGTTGCCACAGAACAAAGACAAGAAGTATACTAACTTGTATAAATTATCTAGGAAAAAAACCT
 TCAGAATTCTAAGATGAATTACCAAGGTGAGAATGAATAAGCTATGCAAGGTATTTGTAATATACTGTGGACAC
 AACCTGCTCTGCCCTCATCTGCCCTAGTGTGCAATCTCATTGACTATACGATAAAAGTTGACAGTCTTACTTT
 CTGTAGAACACTGCCATAGGAATGCTGTTTTTGACTGGACTTTACCTTGATATATGTATATGGATGTATG
 CATAAAATCATAGGACATATGACTTGTGAAACAAGTTGGATTTTATACAATATTAAATTCAACCACCTCAG

FIGURE 15

MEKMLAGCFLLILGQIVLLPAEARERSRGRSISRGRHARTHQPQTALLESSCENKRADLVFII
 DSSRSVNTHDYAKVKEFIVDILQFLDIGPDVTRVGLLQYGSTVKNEFSLKTFKRKSEVERAV
 KMRHLSGTMTGLAIQYALNIAFSEAEGRPLRENVPRVIMIVTDGRPQDSVAEVAAKARD
 TGILIFAIGVGQVDFNTLKSIGSEPHEDHVFLVANFSQIETLTSVFQKKLCTAHMCSTLEHN
 CAHFCINIPGSYVCRCKQGYILNSDQTTCRIQDLCAMEDHNCEQLCVNVPGSFVCQCYSGYA
 LAEDGKRCVAVDYCASENHGCEHECVNADGSYLCQCHEGFALNPDEKTCTRINYCALNKPGC
 EHECVNMEESYYCRCHRGYTLDPNGKTCRVDHCAQQDHGCEQLCLNTEDSFVCQCSEGFLI
 NEDLKTCSRVDYCLLSDHGCEYSCVNMDRSFACQCPEGHVLRSDGKTCAKLDSCALGDHGCE
 HSCVSSEDSFVCQCFCFGYILREDGKTCRKCVCQAIIDHGCEHICVNSDDSYTCECLEGFR
 LAEDGKRCRCKDVCKSTHHGCEHICVNNNGNSYICKCSEGFLVLAEDGRRCKCTEGPIDLVFVID
 GSKSLGEENFEVVVKQFVTGILDSLTISPKAARVGLLQYSTQVHTEFTLRFNSAKDMKKAVA
 HMKYMKGKGSMTGLALKHMFERSFTQGEGARPLSTRVPRAAIVFTDGRAQDDVSEWASKAKAN
 GITMYAVGVGKAIEEELQEIASEPTNKHLFYAEDFSTMDEISEKLKKGICEALEDSDGRQDS
 PAGELPKTVQQPTESEPVTINIQDILSCSFQVQHRYLFEEDNLLRSTQKLSHSTKPGSPL
 EEKHDQCKCENLIMFQNLANEEVRKLTQRLEEMTQRMEALENRLRYR

Signal peptide:

amino acids 1-23

N-glycosylation site.

amino acids 221-225

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 115-119, 606-610, 892-896

Casein kinase II phosphorylation site.amino acids 49-53, 118-122, 149-153, 176-180, 223-227, 243-247,
 401-405, 442-446, 501-505, 624-628, 673-677, 706-710, 780-784,
 781-785, 819-823, 866-870**N-myristoylation site.**amino acids 133-139, 258-264, 299-305, 340-346, 453-459, 494-500,
 639-645, 690-696, 752-758, 792-798**Amidation site.**

amino acids 314-318, 560-564, 601-605

Aspartic acid and asparagine hydroxylation site.amino acids 253-265, 294-306, 335-347, 376-388, 417-423, 458-464,
 540-546, 581-587

FIGURE 16

GGAGCCGCCCTGGGTGTCAGCGGCTGGCTCCCGCGCACGCTCCGGCGTCGCGCAGCCTCG
GCACCTGCAGGTCCGTGCGTCCCGCGGCTGGCGCCCTGACTCCGTCCCGGCCAGGGAGGGC
CATGATTCCCTCCGGGCCCCCTGGTACCAACTTGCTGCGGTTTGTTCCTGGGCTGA
GTGCCCTCGCGCCCCCTCGCGGGCCAGCTGCAACTGCACCTGCCGCCAACCGGTTGCAG
GCGGTGGAGGGAGGGAAAGTGGTCTTCAGCGTGGTACACCTTGACGGGGAGGTGTCTTC
ATCCCAGCCATGGGAGGTGCCCTTGTGATGGTTCTCAAACAGAAAGAAAAGGAGGATC
AGGTGTTGTCCTACATCAATGGGTACAAACAAGCAAACCTGGAGTATCCTGGTCTACTCC
ATGCCCTCCCGAACCTGTCCTGCCGTGGAGGGTCTCCAGGAGAAAGACTCTGGCCCTA
CAGCTGCTCCGTGAATGTGCAAGACAAACAAGGCAAATCTAGGGCCACAGCATCAAACCT
TAGAACTCAATGTACTGGTTCTCAGCTCCATCCTGCCGTCTCCAGGGTGTGCCCAT
GTGGGGCAAACGTGACCCCTGAGCTGCCAGTCTCAAGGAGTAAGCCGCTGTCCAATACCA
GTGGGATCGGCAGCTCCATCCTCCAGACTTCTTGACCCAGCATTAGATGTATCCGTG
AGTGGTTAAGCCTCACCAACCTTCGTCTTCATGGCTGGAGTCTATGTCTGCAAGGCCAC
AATGAGGTGGGACTGCCAATGTAATGTGACGCTGGAAGTGAGCACAGGCCCTGGAGCTGC
AGTGGTTGCTGGAGCTGTTGTGGTACCCCTGGTTGGACTGGGTTGCTGGCTGGCTGGTCC
TCTTGTACCACCGCCGGGCAAGGCCCTGGAGGAGCCAGCCAATGATATCAAGGAGGATGCC
ATTGCTCCCCGGACCCCTGCCCTGCCCAAGAGCTCAGACACAATCTCAAGAAATGGGACCC
TTCCTCTGTACCTCCGCACGAGCCCTCCGCCACCCATGCCCTCCAGGCCCTGGTGCAT
TGACCCCCACGCCAGTCTCTCCAGCCAGGCCCTGCCCTCACCAAGACTGCCACGACAGAT
GGGGCCCAACCAATATCCCCATCCCTGGTGGGTTCTCCTCTGGCTTGAGCCG
CATGGGTGCTGCTGTGATGGTGCCTGCCAGACTCAAGCTGGCTCTGGTATGATGAC
CCCACCACTATTGGCTAAAGGATTGGGCTCTCCTCTATAAGGGTACCTCTAGCAC
AGAGGCCTGAGTCATGGAAAGAGTCACACTCCTGACCCCTAGTACTCTGCCACCCACCTCTC
TTTACTGTGGAAAACCCTCACTCAGTAAGACCTAAGTGTCCAGGAGACAGAAGGAGAAGAGGA
AGTGGATCTGGAATTGGGAGGAGCCTCCACCCACCCCTGACTCCTCCTATGAAGCCAGCTG
CTGAAATTAGCTACTCACCAAGAGTGAGGGCAGAGACTTCCAGTCAGTGAGCTCCAGG
CCCCCTGATCTGTACCCACCCCTATCTAACACCACCCCTGGCTCCACTCCAGCTCCCTGT
ATTGATATAACCTGTCAGGCTGGCTGGTTAGGTTTACTGGGGCAGAGGATAGGGAAATCTC
TTATTAAAACATGAAATATGTGTTTTCTATTGCAAATTAAAGATAACATAA
TGTTGTATGAAAAA

FIGURE 17

MISLPGPLVTNLLRFLFLGLSALAPPSRAQLQLHLPANRLQAVEGGEVVLPWYTLHGEVSS
SQPWEVPFVMWFFKQKEKEDQVLSYINGVTTSKPGVSLVYSMPSRNLSLRLEGLQEKGDSGPY
SCSVNVQDKQGKSRGHSIKTLELNVLVPPAPPSCRLQGVPHGANVTLSCQSPRSKPAVQYQ
WDRQLPSFQTFFAPALDVIRGSLSLTNLSSSMAGVYVCKAHNEVGTACQCNVTLEVSTGPGAA
VVAGAVVGTLVGLGLLAGLVLLYHRRGKALEEPANDIKEDAIAPRTLWPKSSDTISKNGTL
SSVTSARALRPPHGPPRGALTPPSLSSQALPSPRLPTTDGAHPQPISPIPGVSSSGLSR
MGAVPVMVPAQSQAGSLV

Signal peptide:

amino acids 1-29

Transmembrane domain:

amino acids 245-267

N-glycosylation site.

amino acids 108-112, 169-173, 213-217, 236-240, 307-311

N-myristoylation site.

amino acids 90-96, 167-173, 220-226, 231-237, 252-258, 256-262,
262-268, 308-314, 363-369, 364-370

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 164-175

FIGURE 18

CGCCACCACTGCGGCCACCGCCAATGAAACGCCTCCCGCTCCTAGTGGTTTTCCACTTG
 TTGAATTGTTCTATACTCAAAATTGACCAAGACACCTGTCTCCAAATGCAAATGTGA
 AATACGCAATGAAATTGAAGCCTGCTATTGCAACATGGGATTTCAGGAAATGGTGTACAA
 TTTGTGAAGATGATAATGAATGTGAAATTAACTCAGTCTGTGGCGAAATGCTAATTGC
 ACTAACACAGAAGGAAGTTATTATTGTATGTGTACCTGGCTCAGATCCAGCAGTAACCA
 AGACAGGTTATCCTAAATGATGGAACCGTCTGTATAGAAAATGTGAATGCAAACGCCATT
 TAGATAATGTCTGTATAGCTGCAAATATTAAATAAAAACCTTAACAAAATCAGATCCATAAAA
 GAACCTGTGGCTTGCTACAAGAAGTCTATAGAAATTCTGTGACAGATCTTCACCAACAGA
 TATAATTACATATAGAAATTAGCTGAATCATCTTCATTACTAGGTTACAAGAACAAACA
 CTATCTCAGCCAAGGACACCCTTCAACTCAACTCTTACTGAATTGTAAAACCGTGAAT
 AATTTGTTCAAAGGGATACATTGTAGTTGGGACAAGTTATCTGTGAATCATAGGAGAAC
 ACATCTTACAAAACATGCACACTGTTGAAAGCTACTTTAAGGATATCCCAGAGCTTCC
 AAAAGACCACAGAGTTGATACAAATTCAACGGATATAGCTCTCAAAGTTCTTTGAT
 TCATATAACATGAAACATATTCACTCCTCATATGAATATGGATGGAGACTACATAAATATT
 TCCAAGAGAAAAGCTGCATATGATTCAAATGGCAATGTTGCAGTTGCATTGATTTATATTATA
 AGAGTATTGGTCCTTGCTTCACTCATCTGACAACCTTCTATTGAAACCTCAAATTATGAT
 AATTCTGAAGAGGAGGAAAGAGTCATATCTCAGTAATTCTAGTCTCAATGAGCTCAAACCC
 ACCCACATTATGAACCTGAAAAATAACATTACATTAAAGTCATCGAAAGGTACAGATA
 GGTATAGGAGTCTATGTGCTTTGAAATTACTCACCTGATACCATGAATGGCAGCTGGTCT
 TCAGAGGGCTGTGAGCTGACACTCAAATGAGACCCACACCTCATGCCGCTGTAATCACCT
 GACACATTTGCAATTGATGTCCTCTGGCCTTCCATTGGTATTAAAGATTATAATATT
 TTACAAGGATCACTCAACTAGGAATAATTATTCACTGATTGTCTGCCATATGCATT
 ACCTCTGGTCTTCAGTGAATTCAAAGCACCAGGACAACAATTCAACAAAATCTTGCTG
 TAGCCTATTCTGCTGAACTTGTTCTGGATCAATACAATAAGCTCT
 TCTGTTCAATCATGCCGACTGCTACACTACTCTTTAGCTGCTTGCATGGATGTGC
 ATTGAAGGCATACATCTCATCTGTTGTTGGTGTCACTACAACAAGGGATTTGCA
 CAAGAATTATATCTTGGCTATCTAAGCCCAGCCGTGGTAGTTGGATTTCGGCAGCAC
 TAGGATACAGATATTATGGCACAACCAAGTATGGCTTAGCACCAGAAACAACCTTATT
 TGGAGTTTATAGGACCAGCATGCCATTCTGTTAATCTCTGGCTTTGGAGTCAT
 CATATACAAAGTTTCGTCACACTGCAGGGTTGAAACCAGAGTTAGTTGCTTGAGAAC
 TAAGGTCTTGTGCAAGAGGAGCCCTCGCTTCTGTTCTGGCACACCTGGATCTT
 GGGTTCTCCATGTTGTGACGCATCAGTGGTTACAGCTTACCTCTCACAGTCAGCAATGC
 TTTCCAGGGATGTTCTTGTGTTTATCTAGAAAGATTCAAGAAC
 ATTACAGATTGTCAAAATGTCCTGTTGGATGTTAAGGTAACATAGAGAAC
 GTGGATAATTACAACGTGACAAAATTTCAAGCTGTGGATGACCAATGTATAAAA
 TGACTCATCAAATTATCCAATTATTAACACTAGACAAAAAGTATTAAATCAGTTCT
 GTTTATGCTAGGAACGTAGATAATAAGTAAATTATGTATCATATAGATATACTATGT
 TTTCTATGTAATAGTTCTGTCAAAATAGTATTGCAAGATATTGGAAAGTAATTGGTT
 CTCAGGAGTGTATCACTGCACCCAGGAAAGATTCTTCAACACGAGAAGTATATGAA
 TGTCTGAAAGGAAACACTGGCTTGATATTCTGTGACTCGTGTGCTTGAACACTAGTCC
 CCTACCACTCGTAATGAGCTCCATTACAGAAAGTGGAACATAAGAGAAC
 ATATCAAACAGTGAAAGGGAATGATAAGATGTATTGAAATGAACTGTTCTGTAGAC
 TAGCTGAGAAATTGTTGACATAAAATAAGAATTGAAAGAACACATTTACCACTTGTGAA
 TTGTTCTGAACTAAATGTCCACTAAACAACTTAGACTTCTGTTGCTAAATCTGTTCT
 TTTCTAATATTCTAAAAAAAAAGGTTACCTCCACAAATTGAAAAA
 AAAAAAAAAAAAAAAAAAAAAA

FIGURE 19

MKRLPLLVVFSTLLNCSYTQNCTKTPCLPNAKCEIRNGIEACYCNMGFSGNVTICEEDNECGNLTQSCGENANCTNTEGSYYCMCVPGRSSSNQDRFITNDGTVCIENVANCHLDNVIAANINKTLTKIRSIKEPVALLQEYVRNSVTDLSPTDIITYIEILAESSLLGYKNNTISAKDTNSNLTLEFVKTVNPFQRTFVWDKLSVNHRRTHLTKLMHTVEQATLRIQSFSQKTTEFDTNSTDIALKVFFFDSYNMCKIHPHMMNDGDYINIFPKRKAAYDSNGNVAVAFLYYKSIGPLLSSSDNFLLPQNYDNSEEERVISSVISVSMSSNPPTLYELEKITTLSHRKVTDRYRSLCAFWNYSPTDMNGWSSEGCELTYSNETHTSCRCHNLTHFAILMSSGPGSIGIKDYNILTRITQLGIIISLICLAIICIFTFWFSEIQSTRTTIHKNLCCSLFLAELVFLVGINTNTNKLFCIIAGLIIHYFFLAFAWMCIEGIHLYLIVVGVIYNKGFLHKNFYIFGYLSPAVVVGFSAAALGYRYYGTTKVCWLSTENNFIFSFIGPACLIILVNLLAFGVIIYKVFRHTAGLKPEVSCFENIRSCARGALALLFLLGTTWIFGVVLHVHASVVTAYLFTVSNAFQGMFIFLFLCVLSRKIQEEYYRLFKNPCCFGCLR

Signal peptide:

amino acids 1-19

Transmembrane domain:

amino acids 430-450, 465-486, 499-513, 535-549, 573-593, 619-636, 648-664

N-glycosylation site.

amino acids 15-19, 21-25, 64-68, 74-78, 127-131, 177-181, 188-192, 249-253, 381-385, 395-399

Glycosaminoglycan attachment site.

amino acids 49-53

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 360-364

Casein kinase II phosphorylation site.

amino acids 54-58, 68-72, 76-80, 94-98, 135-139, 150-154, 155-159, 161-165, 181-185, 190-194, 244-248, 310-314, 325-329, 346-350, 608-612

Tyrosine kinase phosphorylation site.

amino acids 36-44, 669-677, 670-678

N-myristoylation site.

amino acids 38-44, 50-56, 52-58, 80-86, 382-388, 388-394, 434-440, 480-486, 521-527

Aspartic acid and asparagine hydroxylation site.

amino acids 75-87

FIGURE 20

TGGAAACATATCCTCCCTCATATGAATATGGATGGAGACTACATAAATATTTCAAAGNG
AAAAGCCGGCATATGGATTCAAATGCAATGTTGCAGTTGCATTTTATATTATAAGAGTAT
TGGTCCCTTGCTTCATCATCTGACAACCTCTTATTGAAACCTCAAAATTATGATAATTCT
GAAGAGGAGGAAAGAGTCATATCTTCAGTAATTCAGTCTCAATGAGCTCAAACCCACCCAC
ATTATATGAACTTGAAAAAAACATTTACATTAAGTCATCGAAAGGTACAGATAGGTATA
GGAGTCTATGTGGCATTGGAATACTCACCTGATACCATGAATGGCAGCTGGTCTTCAGAG
GGCTGTGAGCTGACACTCAAATGAGACCCACACCTCATGCCGCTGTAATCACCTGACACA
TTTGCAATTTGATGTCCTCTGGCTTCCATTGGTATTAAAGATTATAATATTCTTACAA
GGATCACTCAACTAGGAATAATTATTCACTGATTGTCTGCCATATGCATTTTACCTTC
TGGTTCTTCAGTGAATTCAAAGCACCAGGA

FIGURE 21

GCTCCCAGCCAAGAACCTCGGGGCCGCTGCGCGGTGGGGAGGAGTTCCCCGAAACCCGGCCG
CTAAGCGAGGCCTCCCTCCCGCAGATCCGAACGGCCTGGGCGGGGTACCCCGGCTGGGA
CAAGAACGCCGCCCTGCCTGCCCGGGCCGGAGGGGGCTGGGGCTGGGGCCGGAGGCAGG
GGTGTGAGTGGGTGTGTGCGGGGGCGGAGGCTTGATGCAATCCCGATAAGAAATGCTCGGG
TGTCTGGGACCTACCCGTGGGCCCCGTAAGGCGCTACTATATAAGGCTGCCGGCCCGAG
CCGCCGCGCCGTAGAGCAGGAGCGCTCGCTCCAGGATCTAGGGCACGACCATCCAACCC
GGCACTCACAGCCCCGAGCGCATCCCGTCGCCGCCAGCCTCCGCACCCCCATGCCGG
AGCTGCCCGAGAGCCCCAGGGAGGTGCCATCGGAGCGGGTGTGTGGTGTCCACGTATGG
ATCCTGGCCGGCCTCTGGCTGGCGTGGCGGGCGCCCCCTGCCCTCTCGGACGCCGG
CCACGTGCACTACGGCTGGGGGACCCCATCCGCCTCGGGCACCTGTACACCTCCGGCCCC
ACGGGCTCTCCAGCTGCTCCCTGCGCATCCGTGCCGACGGCGTGTGGACTGCGCGGGGC
CAGAGCGCGCACAGTTGCTGGAGATCAAGGCAGTCGCTCTCGGACCGTGGCCATCAAGGG
CGTGCACAGCGTGCCTGACCTCTGCATGGGCCGACGGCAAGATGCAAGGGCTGCTTCAGT
ACTCGGAGGAAGACTGTGCTTCGAGGAGGAGATCCGCCAGATGGCTACAATGTGTACCGA
TCCGAGAAGCACCGCCTCCGGTCTCCCTGAGCAGTGCACAGCGGAGCTGTACAAGAA
CAGAGGCTTCTCCACTCTCTCATTTCTGCCATGCTGCCATGGTCCCAGAGGAGCCTG
AGGACCTCAGGGGCCACTTGAATCTGACATGTTCTCTCGCCCTGGAGACCGACAGCATG
GACCCATTGGGCTTGTCAACCGACTGGAGGCCGTGAGGAGTCCAGCTTGAGAAGTAAC
GAGACCATGCCCGGGCCTTCACTGCTGCCAGGGCTGTGGTACCTGCAGCGTGGGGACG
TGCTTCTACAAGAACAGTCTGAGTCCACGTTCTGTTAGCTTAGGAAGAAACATCTAGAA
GTTGTACATATTCAAGAGTTCCATTGGCAGTGCCTAGTTCTAGCCAATAGACTTGTCTGAT
CATAACATTGTAAGCCTGTAGCTTGCCTGCCAGCTGCTGCCCTGGGCCATTCTGCTCCCTCGA
GGTTGCTGGACAAGCTGCTGCACTGCTCAGTTCTGCTGAATACCTCCATCGATGGGAAC
TCACTCCCTTGGAAAAATTCTTATGTCAAGCTGAAATTCTCTAATTTCATCACTTC
CCCAGGAGCAGCCAGAACAGACAGGCAGTAGTTTAATTCAAGGAACAGGTGATCCACTCTGTA
AAACAGCAGGTAAATTCACTCAACCCCATGGGAATTGATCTATATCTACTTCCAGGG
ACCATTGCCCTCCCAAATCCCTCCAGGCCAGAACTGACTGGAGCAGGATGCCACCAG
GCTTCAGGAGTAGGGGAAGCCTGGAGCCCCACTCCAGCCCTGGGACAACCTGAGAATTCCCC
CTGAGGCCAGTTCTGTCACTGGATGCTGTCTGAGAATAACTTGCTGTCCGGTGTACCTGC
TTCCATCTCCAGCCCACCAGCCCTGCCCACCTCACATGCCTCCCCATGGATTGGGCCT
CCCAGGCCCCCACCCTTATGTCAACCTGCACTTCTGTTCAAAATCAGGAAAAGAAAGAT
TTGAAGACCCCAAGTCTGTCAATAACTTGCTGTGGAAAGCAGCGGGGAAGACCTAGAAC
CCTTCCCCAGCACTTGGTTTCCAACATGATATTATGAGTAATTATTGATATGTACA
TCTCTTATTTCATTACATTATTATGCCCTAAATTATATTATGTATGTAAAGTGGAGTTGTTGT
TTTGTATATTAAATGGAGTTGTTGT

FIGURE 22

MRSGCVVHVWILAGLWLAVAGRPLAFSDAGPHVHYGWGDPIRLRHLYTSGPHGLSSCFLRI
RADGVVDCARGQSAHSLLEIKAVALRTVAIKGVHSVRYLCMGADGKMQGLLQYSEEDCAFEE
EIRPDGYNVYRSEKHRLPVSLSSAKQRQLYKNRGFLPLSHFLPMLPMVPEEPEDLRGHLESD
MFSSPLETDSMDPFGLVTGLEAVRSPSFEK

Signal peptide:

amino acids 1-22

Casein kinase II phosphorylation site.

amino acids 78-82, 116-120, 190-194, 204-208

N-myristoylation site.

amino acids 15-21, 54-60, 66-72, 201-207

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 48-59

FIGURE 23

CCCAGAACGTTCAAGGGCCCCGGCCTCCTGCGCTCCTGCCGCCGGGACCCCTGACCTCCTCA
GAGCAGCCGGCTGCCGCCGGGAAGATGGCGAGGAGGAGCCACCGCCTCCTGCTG
CTGCTGCGCTACCTGGTGGTCGCCCTGGCTATCATAAGGCCTATGGGTTTCTGCCCAAA
AGACCAACAAGTAGTCACAGCAGTAGAGTACCAAGAGGCTATTTAGCCTGCAAACCCCAA
AGAAGACTGTTCTCCAGATTAGAGTGGAAAGAACTGGGTGGAGTGTCTCCTTGTCTAC
TATCAACAGACTCTTCAAGGTGATTTAAAAATCGAGCTGAGATGATAGATTCAATATCCG
GATCAAAAATGTGACAAGAAGTGATGCCGGAAATATCGTTGTGAAGTTAGTGGCTCCAGCAGTT
AGCAAGGCCAAACCTGGAAGAGGATACAGTCACTCTGGAAGTATTAGTGGCTCCAGCAGTT
CCATCATGTGAAGTACCCCTTCTGCTCTGAGTGGAACTGTGGTAGAGCTACGATGTCAAGA
CAAAGAAGGGAAATCCAGCTCCTGAATACACATGGTTAAGGATGGCATCCGTTGCTAGAAA
ATCCCAGACTGGCTCCCAAAGCACCAACAGCTCATACACAATGAATAACAAAAACTGGAAC
CTGCAATTAAATACTGTTCCAAACTGGACACTGGAGAATATTCCGTGAAGCCGCAATT
TGTGGATATCGCAGGTGTCTGGAAACGAATGCAAGTAGATGATCTAACATAAGTGGCA
TCATAGCAGCCGTAGTAGTTGTGGCCTAGTGATTCCGTTGTGGCCTTGGTGTATGCTAT
GCTCAGAGGAAAGGCTACTTTCAAAAGAAACCTCCTCCAGAAGAGTAATTCTCATCTAA
AGCCACGACAATGAGTGAAGATGTGAGGTCACTGGCTCACGCCGTAACTCCAGCACTTGGAGG
CCGGCGGGCGGATCACGAGGTCAAGGAGTTCTAGACCAGTCTGGCAATATGGTAAACCC
CATCTCTACTAAAATACAAAAATTAGCTGGCATGGTGGCATGTGCCTGCAGTCCAGCTGC
TTGGGAGACAGGAGAATCACTGAACCCGGGAGGGCGGAGGTTGCAGTGAGCTGAGATCACGC
CACTGCAGTCCAGCCTGGTAACAGAGCAAGATTCCATCTAAAAAATAAAATAATA
AATAAAATACTGGTTTTACCTGTAGAATTCTTACAATAATAGCTTGATATT

FIGURE 24

MARRSRHRLLLLLRYLVVALGYHKAYGFSAPKQQVVTAVEYQEAILACKTPKKTVSSRLE
WKKLGRSVSFVYYQQLQGDFKNRAEMIDFNIRIKNVTRSDAGKYRCEVSAPSEQGQNLEED
TVTLEVLVAPAVPSCEVPSSALSGTVVELRCQDKEGNPAPEYTWFKDGI
LLENPRLGSQST
NSSYTMTKTGTLQFNTVSKLDTGEYSCEARN
SVGYRRCPGKRMQVDDLNISGII
AAVVVVA
LVISVCGLGVCYAQRKGYFSKETSFQKSNSSSKATTMSENVQWLTPV
PALWKA
AAAGGSRGQEF

Signal peptide:

amino acids 1-20

Transmembrane domain:

amino acids 130-144, 238-258

N-glycosylation site.

amino acids 98-102, 187-191, 236-240, 277-281

Casein kinase II phosphorylation site.

amino acids 39-43, 59-63, 100-104, 149-153, 205-209, 284-288

N-myristoylation site.

amino acids 182-188, 239-245, 255-261, 257-263, 305-311

Amidation site.

amino acids 226-230

FIGURE 25

GACATCGGAGGTGGCTAGCACTGAAACTGCTTTCAAGACGAGGAAGAGGGAGGAGAAAGAG
 AAAGAAGAGGAAGATGTTGGCAACATTTATTAACATGCTCCACAGCCGGACCTGGCAT
 CATGCTGCTATTCTGCAAATACTGAAGAAGCATGGGATTAAATATTTACTTCTAAATAA
 ATGAATTACTCAATCTCCTATGACCACATCTACATACACTCCACCTCAAAAAGTACATCAATA
 TTATATCATTAAGGAAATAGTAACCTCTCTCCAAATATGCATGACATTGGACAATG
 CAATTGTGGCACTGGCACTTATTCAGTGAAGAAAAACTTGTGGTTCTATGGCATTCA
 TTTGACAAATGCAAGCATTCCCTTATCAATCAGCTCCTATTGAACCTACTAGCACTGACTG
 TGGAACTCTTAAGGGCCCATTACATTCTGAAGAAGAAAGCTAAGATGAAGGACATGCCACT
 CGAATTCACTGTGCTACTTGGCTAGCTATCACTACACTAGTACAAGCTGTAGATAAAAAAG
 TGGATTGTCCACGGTTATGTACGTGTGAAATCAGGCCTGGTTACACCCAGATCCATTAT
 ATGGAAGCATCTACAGTGGATTGTAATGATTAGGTCTTTAACTTCCCAGCCAGATTGCC
 AGCTAACACACAGATTCTCTCCTACAGACTAACAAATATTGCAAAATTGAATACTCCACAG
 ACTTTCAGTAAACCTTACTGGCCTGGATTATCTCAAAACAATTATCTCAGTCACCAAT
 ATTAATGAAAAAGATGCCTCAGCTCCTTCTGTGTACCTAGAGGAAAACAAACTTACTGA
 ACTGCCGAAAAATGTCTGTCGAACGTGAGCAACTACAAGAACTCTATATTAAATCACAAC
 TGCTTCTACAATTTCACCTGGAGCCTTATTGGCCTACATAATCTCTCGACTCATCTC
 AATTCAAATAGATTGCAGATGATCAACAGTAAGTGGTTGATGCTCTCCAAATCTAGAGAT
 TCTGATGATTGGGAAAATCCAATTATCAGAACAGACATGAACTTTAAGCCTTTATCA
 ATCTTCAGCCTGGTTATAGCTGGTATAAACCTCACAGAAATACAGATAACGCCTGGTT
 GGACTGGAAAACCTAGAAAGCATCTCTTTACGATAACAGGCTTATTAAAGTACCCATGT
 TGCTCTCAAAAAGTTGTAATCTCAAATTGGATCTAAATAAAATCCTATTAAATAGAA
 TACGAAGGGGTGATTTAGCAATATGCTACACTAAAAGAGTTGGGATAAATAATATGCCT
 GAGCTGATTTCCATCGATAGTCTGCTGTGGATAACCTGCCAGATTTAAGAAAAATAGAAGC
 TACTAACACCCCTAGATTGTCTACATTACCCCCAATGCATTTCAGACTCCCCAAGCTGG
 AATCACTCATGTCGAACAGCAATGCTCTCAGTGCCTGTACCATGGTACCATGGTAC
 CCAAAACCTCAAGGAAATCAGCATACACAGTAACCCCATCAGGTGTGACTGTGT
 GATGAACATGAACAAAACCAACATTGATTGACGGACATTCACTGTTGGTGGACC
 CACCTGAATTCCAAGGTCAGAATGTCGGCAAGTGCATTCAAGGACATGATGGAAATTGT
 CTCCTCTTATAGCTCTGAGAGCTTCCTTAATCTAAATGTAGAAGCTGGAGCTATGT
 TTCCTTCAGTGTAGAGCTACTGCAGAACACAGCCTGAAATCTACTGGATAACACCTCTG
 GTCAAAAACCTTGCTTAATACCCCTGACAGAACAGTTCTATGTCCATTCTGAGGGAAACACTA
 GATATAAAATGGCTTAACCTCCAAAGAAGGGGGTTATATACTTGTATAGCAACTAACCTAGT
 TGGCGCTGACTTGAAGTCTGTTATGATCAAAGTGGATGGATCTTCCACAAGATAACAATG
 GCTCTTGAATTAAAATAAGAGATATTCAAGGCCATTCAAGTTGGTGTCTGGAAAGCA
 AGTTCTAAAATTCTCAAATCTAGTGTAAATGGACAGCCTTGTCAAGACTGAAAATTCTCA
 TGCTGCGCAAAGTGCCTGAATACCATCTGATGTCAAGGTATATACTTACTCATCTGAATC
 CATCAACTGAGTATAAAATTGTATTGATATTCCACCATCTACAGAAAAACAGAAAAAAA
 TGTGTAAATGTCAACCACCAAGGTTGCACCTGATCAAAAGAGTATGAAAAGATAATAC
 CACAAACACTTATGGCCTGTCTGGAGGCCTCTGGGGATTATTGGTGTGATATGTCTTATCA
 GCTGCCCTCTCCAGAAATGAACTGTGATGGGGACACAGCTATGTGAGGAATTACTACAG
 AAACCAACCTTGCATTAGGTGAGCTTATCTCCTCTGATAAAATCTCTGGGAAGCAGGAAA
 AGAAAAAAAGTACATCACTGAAAGTAAAGCAACTGTTAGGTTACCAACAAATATGTCCT
AAAAACCAAGGAAACCTACTCCAAAATGAAC

FIGURE 26

MKDMPLRIHVLLGLAITTLVQAVDKVDCPRLCTCEIRPWFTPRS IYMEASTVDCNDLQLLT
FPARLPANTQILLQTNNIAKIEYSTDGPVNLTGQLDLSQNNLSSVTNINVKKMPQLLSVYLE
ENKLTELPEKCLSELSNLQELYINHNLLSTISPGAFIGLHNLLRLHLSNRLQMINSKWFDA
LPNLEILMIGENPIIRIKDMNFKPLINRLRSLVIAGINLTEIPDNAVGLENLESISFYDNRL
IKVPHVALQKVVNLIKFLDLNKNPINRIRRGDFSNMLHLKELGINNMPELISIDS LAVDNLPD
LRKIEATNNPRLSYIHPNAFFRLPKLESMLNSNALSALYHGTIESLPNLKEISIHSNPIRC
DCVIRWMNMNKTNIRFMEPDSDLFCVDPPEFQGQNVROVHFRDMMEICLPLIAPESFPSNLNV
EAGSYVFSFHCRATAEPQPEIYWITPSGQKLLPNTLTDKFYVHSEGTLDINGVTPKEGGLYTC
IATNLVGDALKSVMIKVDGSFPQDNNGSLNIKIRDIQANSVLVSWKASSKILKSSVKWTAFV
KTENSHAAQSARI PSDVKVYNLTHLN PSTEYKICIDIPTIYQKNRKKCVNVT KGLHPDQKE
YEKNNTTLMACLGLLGIIGVICLISCLSPEMNC DGGHSYVRN YLQKPTFALGELYPLIN
LWEAGKEKSTS LKV KATV IGLPTNMS.

Signal sequence:

amino acids 1-22

Transmembrane domain:

amino acids 633-650

N-glycosylation site.

amino acids 93-97, 103-107, 223-227, 382-386, 522-526, 579-583,
608-612, 624-628, 625-629

Casein kinase II phosphorylation site.

amino acids 51-55, 95-99, 242-246, 468-472, 487-491

Tyrosine kinase phosphorylation site.

amino acids 570-579

N-myristoylation site.

amino acids 13-19, 96-102, 158-164, 221-227, 352-358, 437-443,
491-497, 492-498, 634-640, 702-708

Cell attachment sequence.

amino acids 277-280

FIGURE 27

GCCCCGGGACTGGCGCAAGGTGCCAAGCAAGGAAAGAAATAATGAAGAGACACATGTGTTAG
CTGCAGCCTTTGAAACACGCAAGAAGGAAATCAATAGTGTGGACAGGGCTGGAACCTTTAC
CACGCTTGGAGTAGATGAGGAATGGGCTCGTATTGCTGACATTCCAGCATGAATCT
GGTAGACCTGTGGTTAACCGTTCCCTCTCCATGTGTCTCCTACAAAGTTGTTCTTA
TGATACTGTGCTTCATTCTGCCAGTATGTGTCCAAGGGCTGTCTTGTCTCCTCTGGG
GGTTAAATGTCACCTGTAGCAATGCAAATCTCAAGGAAATACCTAGAGATCTCCTCCTGA
AACAGTCTTACTGTATCTGGACTCCAATCAGATCACATCTATTCCAATGAAATTAAAGG
ACCTCCATCAACTGAGAGTTCTCAACCTGTCCAAAATGGCATTGAGTTATCGATGAGCATT
GCCTTCAAAGGAGTAGCTGAAACCTTGCAAGACTCTGGACTTGTCCGACAATCGGATTCAAAG
TGTGCACAAAATGCCTCAATAACCTGAAGGCCAGGGCCAGAATTGCCAACACCCTGGC
ACTGCGACTGTACTCTACAGCAAGTCTGAGGAGCATGGCGTCCAATCATGAGACAGCCCAC
AACGTGATCTGTAACACGTCCGTGTGGATGAACATGCTGGCAGACCATTCCCTCAATGCTGC
CAACGACGCTGACCTTGTAAACCTCCCTAAAAAAACTACCGATTATGCCATGCTGGTCACCA
TGTGGCTGGTTCACTATGGTGTCTCATATGTGGTATATTATGTGAGGCAAATCAGGAG
GATGCCGGAGACACCTCGAATACTTGAAATCCCTGCCAAGCAGGCAGAAGAAAGCAGATGA
ACCTGATGATATTAGCACTGTGGTATAGTGTCCAAACTGACTGTCATTGAGAAAGAAAGAAA
GTAGTTGCGATTGCAGTAGAAATAAGTGGTTACTTCTCCATCCATTGTAACACATTGAA
ACTTTGTATTCAGTTTTGAATTATGCCACTGCTGAACCTTAACAAACACTACAACA
TAAATAATTGAGTTAGGTGATCCACCCCTTAATTGTACCCCCGATGGTATATTCTGAGT
AAGCTACTATCTGAACATTAGTTAGATCCATCTCACTATTAAATAATGAAATTATTTTT
AATTAAAAGCAAATAAAAGCTTAACCTTGAAACCATGGGAAAAAAAAAAAAAAACA

FIGURE 28

MNLVDLWLTRSLSMCLLQSFVLMILCFHSASMC PKGCLCSSGGLNVTCSNANLKEIPRDL
PPETVLLYLDSNQITSIPNEIFKDLHQLRVLNLSKNGIEFIDEHAFKGVAETLQTLSDLSDNR
IQSVHKNAFNNLKARARIANNPWHCDCTLQQVLRSMASNHETAHNVICKTSVLDEHAGRPFL
NAANDADLCNLPKKTTDYAMLVTMFGWFTMVISYVYYYVRQNQEDARRHLEYLKSLPSRQKK
ADEPDDISTVV

Signal sequence:

amino acids 1-33

Transmembrane domain:

amino acids 205-220

N-glycosylation site.

amino acids 47-51, 94-98

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 199-203

Casein kinase II phosphorylation site.

amino acids 162-166, 175-179

N-myristoylation site.

amino acids 37-43, 45-51, 110-116

FIGURE 29

ACCGAGCCGAGCGGACCGAAGGCGGCCGAGATGCAGGTGAGCAAGAGGATGCTGGGGGG
 GCGTGAGGAGCATGCCAGCCCCCTCCTGGCCTGCTGGCAGCCATCCTCTGCTGGTGCT
 GGGCTCAGTGTGTCAGGCTCGGCCACGGCTGCCGCCCCGCTGCGAGTGTCCGCCAGG
 ACCGCCTGTGCTGTGCCACCGCAAGTGCTTGAGCTCCCGAGGGCATCCCCACCGAG
 ACGGCCTGCTGGACCTAGGCAAGAACCGCATAAAACGCTAACCGAGGACGGAGTTGCCAG
 CTTCCCGCACCTGGAGGAGCTGGAGCTCAACGAGAACATCGTGAGGCCGTGGAGCCCG
 CCTTCACAAACCTCTCACCTCCGGACGCTGGGTCTCCGAGCAACGCCCTGAAGCTCATC
 CCGCTAGGCGTCTCACTGCCCTCAGAACCTGACCAAGCAGGACATCAGCGAGAACAGAT
 CGTTATCCTACTGGACTACATGTTAGGACCTGTACAACCTCAAGTCACTGGAGGTTGGCG
 ACAATGACCTCGTCTACATCTCACCGCGCCTCAGGGCCTAACAGCCTGGAGCAGCTG
 ACGCTGGAGAAATGCAACCTGACCTCCATCCCCACCGAGGCCGTGCCCCACCTGCACGCC
 CATCGTCTGAGGCTCCGGCACCTAACATCAATGCCATCCGGACTACTCCTCAAGAGGC
 TGTACCGACTCAAGGTCTGGAGATCTCCACTGGCCTACTGGACACCATGACACCCAAC
 TGCCTCTACGCCCTAACCTGACGTCCCTGTCCATCACACACTGCAATCTGACCGCTGTGCC
 CTACCTGGCGTCCGCCACCTAGTCTATCTCCGTTCTAACCTCTCCTACAACCCATCA
 GCACCATTGAGGCTCCATGTTGATGAGCTGCTCCGGCTGCAGGAGATCCAGCTGGTGGC
 GGGCAGCTGGCCGTGGAGGCCATGCTCCCGCCCTAACACTACCTGCGCGTGTCAA
 TGTCTCTGGCAACCAGCTGACCAACTGGAGGAATCAGTCTTCACTCGGTGGCAACCTGG
 AGACACTCATCTGGACTCCAACCCGCTGGCCTGCGACTGTGCGCTCCTGTGGTGTCCGG
 CGCCGCTGGCGGCTCAACTCAACCGGAGCAGGCCACGTGCGCACGCCAGTTGTCCA
 GGGCAAGGAGTTCAAGGACTTCCCTGATGTGCTACTGCCCAACTACTTCACCTGCCGC
 CCCGCATCCGGGACCGCAAGGCCAGCAGGTGTTGAGGCCACACGGTGCAGTT
 GTGTGCCGGGCCATGGCGACCCGCCCGCATCCTCTGGCTCTCACCCGAAAGCACCT
 GGTCTCAGCCAAGAGCAATGGCGGCTCACAGTCTCCCTGATGGCACGCTGGAGGTGCGCT
 ACGCCCAGGTACAGGACAACGGCACGTACCTGTGATCGCGGCCAACGCGGGCGAACGAC
 TCCATGCCGCCACCTGCATGTGCGCAGCTACTGCCGACTGGCCCCATAGCCAAACAA
 GACCTTCGCTTCATCTCCAACCAGCGGGCGAGGGAGAGGCCAACAGCACCGCGCACTG
 TGCCTTCCCTTCGACATCAAGACCTCATCATGCCACCCATGGCTCATCTCTTTC
 CTGGCGCTGCTCCTCTGCTGGCTGCTGTTCTCTGGAGGCCGGCAAGGGCAACAC
 AAAGCACAACATCGAGATCGAGTATGTGCCCCAAAGTCGGACGCAGGCATCAGCTCCGCC
 ACGCGCCCCGCAAGTTCAACATGAAGATGATATGAGGCCGGGGCGGGGGCAGGGACCCCG
 GGCAGGCCGGCAGGGGAAGGGGCTGGCGCCACCTGCTCACTCTCCAGTCCTTCCACCTC
 CTCCCTACCCCTACACACGTTCTTTCTCCCTCCGCCCTCCGTCCCCCTGCTGCCCG
 CCAGCCCTACCCACCTGCCCTCTTCTACCGAGACCTCAGAACGCCAGACCTGGGACCCCA
 CCTACACAGGGGATTGACAGACTGGAGTTGAAAGCCGACGAACGGACACGCCAGAGTC
 ATAATTCAATAAAAAGTTACGAACCTTCTGTAACTTGGTTCAATAATTATGGATTT
 TATGAAAACCTGAAATAATAAAAAGAGAAAAAACTAAAAAAACAAAAAAACAAAAAA

FIGURE 30

MQVSKRMLAGGVRSMPSPLLACWQPILLVLGSVLSGSATGCPPRCECSAQDRAVLCHRKCF
VAVPEGIPTETRLLDLGKNRIKTLNQDEFASFPHLEELENENIVSAVEPGAFNNLFNLRTL
GLRSNRLKLIPLGVFTGLSNTKQDISENKIVILLDYMFDQDLYNLKSLEVGDNLDLVYISHRA
FSGLNSLEQLTLEKCNLTSIPTEALSHLHGLIVLRLRHLNINAIRDYSFKRLYRLKVLEISH
WPYLDTMTPNCLYGLNLTSLSITHCNLTAVPYLAVRHLVYLRFLNLSYNPISTIEGSMLHEL
LRLQEIQLVGGQLAVVEPYAFRGLNYLRVNVSGNQLTTLEESVFHSVGNLETLILDSNPLA
CDCRLLWVFRRRWRLNFNRQOPTCATPEFVQGKEFKDFPDVLLPNYFTCRRARIRDRKAQQV
FVDEGHTVQFVCRADGDPPIALWLSPRKHLVSAKSNGRLTVFPDGTLEVRYAQVQDNGTYL
CIAANAGGNDSMPAHLHVRSYSPDWPHQPNKTFAFISNQPGEGEANSTRATVPFPFDIKTLI
IATTMGFISFLGVVLFCVLFLWSRGKGNTKHNIEEYVPRKSDAGISSADAPRKFNMKMI

Signal sequence:

amino acids 1-41

Transmembrane domain:

amino acids 556-578

N-glycosylation site.

amino acids 144-148, 202-206, 264-268, 274-278, 293-297, 341-345,
492-496, 505-509, 526-530, 542-546

Casein kinase II phosphorylation site.

amino acids 49-53, 108-112, 146-150, 300-304, 348-352, 349-353,
607-611

Tyrosine kinase phosphorylation site.

amino acids 590-598

N-myristoylation site.

amino acids 10-16, 32-38, 37-43, 113-119, 125-131, 137-143,
262-268, 320-326, 344-350, 359-365, 493-499, 503-509, 605-611

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 32-43

FIGURE 31

CCACACGGTCCGACCTCGCCCCGGCTCCGAAGCGGCTGGGGCGCCCTTCGGTCAAC
ATCGTAGTCCACCCCTCCCCATCCCCAGCCCCGGGATTCAAGGCTGCCAGCGCCAGCC
AGGGAGCCGGCCGGGAAGCGCGATGGGGCCCCAGCCGCCTCGCTCCTGCTCCTGC
TGTTCGCCTGCTGCTGGCGCCGGCGGGCCAACCTCTCCAGGACAGCCAGCCCTGG
ACATCTGATGAAACAGTGGTGGCTGGTGGCACCGTGGTCAAGTGCCAAGTGAAAGATCA
CGAGGACTCATCCCTGCAATGGTCTAACCTGCTCAGCAGACTCTACTTTGGGAGAAGA
GAGCCCTCGAGATAATCGAATTCACTGGTTACCTCTACGCCAACGAGCTCAGCATCAGC
ATCAGCAATGTGCCCTGGCAGACGAGGGCAGTACACCTGCTCAATCTCACTATGCTGT
GCGAACTGCCAAGTCCCTCGTCACTGTGCTAGGAATTCCACAGAACGCCATCATCACTGGTT
ATAAAATCTTCATTACGGGAAAAAGACACAGCCACCCCTAAACTGTCACTTCTGGGAGCAAG
CCTGCAGCCGGCTCACCTGGAGAAAGGGTACCAAGAACACTCCACGGAGAACCAACCCGCAT
ACAGGAAGATCCAATGGTAAAACCTCACTGTCACTGGTACAGCTCGGTGACATTCCAGGTTACCC
GGGAGGATGATGGGGCGAGCATCGTGTGCTGTGAACCATGAATCTCTAAAGGGAGCTGAC
AGATCCACCTCTCAACGCATTGAAGTTTATACACACCAACTGCGATGATTAGGCCAGACCC
TCCCCATCCTCGTGGGCCAGAACAGCTGTTGCTACACTGTGAGGGTGCAGGCAATCCAGTCC
CCCAGCAGTACCTATGGGAGAAGGAGGGCAGTGTGCCACCCCTGAAGATGACCCAGGAGGT
GCCCTGATCTTCCCTTCCCAACAAGAGTGACAGTGGCACCTACGGCTGCAGCCACCA
CAACATGGGAGCTACAAGGCCTACTACACCCCTCAATGTTAATGACCCAGTCCGGTGCCT
CCTCCTCCAGCACCTACCAGCCATCGTGGGATCGTGGCTTCATTGTCTTCTGCTG
CTCATCATGCTCATCTTCTTGGCACTACTTGATCCGGCACAAAGGAACCTACCTGACACA
TGAGGCAAAAGGCTCCGACGATGCTCCAGACGGGACACGGCCATCATCAATGCGAGAAGGCG
GGCAGTCAGGAGGGGACGACAAGAAGGAATATTTCATTAGAGGCGCTGCCACTTCC
GCCCCCCAGGGGCCCTGTGGGACTGCTGGGCCGTACCAACCCGGACTTGTACAGAGCAA
CCGCAGGGCCGCCCTCCGCTTGCTCCAGCCCACCCACCCCTGTACAGAATGTCTGC
TTGGGTGCGGTTGTACTCGTTGGAATGGGAGGGAGGGAGGGGGAGGGAGGG
TTGCCCTCAGCCCTTCCGTGGCTCTGCATTGGTTATTATTATTTGTAACAATCC
CAAATCAAATCTGTCTCCAGGCTGGAGAGGCAGGAGCCCTGGGTGAGAAAAGCAAAAACA
AACAAAAAACA

FIGURE 32

MGAPAASLLLLLFLACCWAPGGANLSQDDSQWPWTSDETVVAGGTVVLKCQVKDHEDSLQW
SNPAQQTLYFGEKRALRDNRIQLVTSTPHELYSISISNVALADEGEYTCSIFTMPVRTAKSLV
TVLGIPQKPIITGYKSSLREKDTATLNCQSSGSKPAARLTWRKGDQELHGEPTRIQEDPNGK
TFTVSSSVTFQVTREDDGASIVCSVNHESLKGADRSTSQRIEVLYTPTAMIRPDPPHPREGQ
KLLLHCEGRGNPVPQQYLWEKEGSVPLKMTQESALIFFPFLNKSDSGTYGCTATSNMGSYKA
YYTLNVNDPSPVPSSSSTYHAIIGGIVAFIVFLLLIMLIFLGHYLIRHKGTYLTHEAKGSDD
APDADTAIINAEGGQSGGDDKKEYFI

Signal sequence:

amino acids 1-20

Transmembrane domain:

amino acids 331-352

N-glycosylation site.

amino acids 25-29, 290-294

Casein kinase II phosphorylation site.

amino acids 27-31, 35-39, 89-93, 141-145, 199-203, 388-392

N-myristoylation site.amino acids 2-8, 23-29, 156-162, 218-224, 295-301, 298-304,
306-310, 334-340, 360-364, 385-389, 386-390**Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 7-18

FIGURE 33

GGGGGTTAGGGAGGAAGGAATCCACCCCCACCCCCCAAACCCCTTCTCCTTCTGG
 CTTCGGACATTGGAGCACTAAATGAACCTGAATTGTGTCGTGGCGAGCAGGATGGTCGCTG
 TTACTTTGTGATGAGATCGGGGATGAATTGCTCGCTTAAAAATGCTGCTTGGATTCTGTT
 GCTGGAGACGTCTTGTGTTGCCGCTGGAAACGTTACAGGGGACGTTGCAAAGAGAAGA
 TCTGTTCTGCAATGAGATAGAAGGGGACCTACACGTAGACTGTGAAAAAAAGGGCTTCACA
 AGTCTGCAGCGTTCACTGCCCGACTTCCAGTTACCATTTCTGCATGGCAATT
 CCTCACTCGACTTTCCCTAATGAGTCGTAACCTTATAATGCGGTTAGTTGCACATGG
 AAAACAAATGGCTTGCATGAAATGTTCCGGGGCTTCTGGGGCTGCAGCTGGTAAAAGG
 CTGCACATCAACAACAAGATCAAGTCTTCGAAAGCAGACTTTCTGGGGCTGGACGA
 TCTGGAATATCTCAGGCTGATTTAATTACGAGATATAGACCCGGGGCCTCCAGG
 ACTTGAAACAAGCTGGAGGTGCTCATTTAAATGACAATCTCATCAGCACCCCTACCTGCCAAC
 GTGTTCCAGTATGTGCCATACCCACCTCGACCTCCGGGTAACAGGCTGAAAACGCTGCC
 CTATGAGGAGGTCTGGAGCAAATCCCTGGTATTGCGGAGATCCTGCTAGAGGATAACCTT
 GGGACTGCACCTGTGATCTGCTCTCCCTGAAAGAATGGCTGAAAACATTCCAAGAATGCC
 CTGATCGGCCAGTGGTCTGCAAGCCCCCACCAGACTGCAGGGTAAAGACCTCAATGAAAC
 CACCGAACAGGACTTGTGTCCTTGGAAAACCGAGTGGATTCTAGTCTCCGGCGCCCCCTG
 CCCAAGAACAGAACCTTGCTCCTGGACCCCTGCCAACTCCTTCAAGACAAATGGCAAGAG
 GATCATGCCACACCAGGGCTGCTCAAACGGAGGTACAAAGATCCAGGAACTGGCAGAT
 CAAAATCAGACCCACAGCAGCGATAGCAGCGGGTAGCTCCAGGAACAAACCCCTAGCTAAC
 GTTTACCTGCCCCGGGGCTGAGCTGCGACCACATCCCAGGGTGGTTAAAGATGAAC
 TGCAACAAACAGGAACGTGAGCAGCTGGCTGATTGAAGCCCAAGCTCTAACGTGCAGGA
 GCTTTCTACGAGATAACAAGATCCACAGCATCCGAAAATCGACTTGTGGATTACAGA
 ACCTCATTCTGTTGGATCTGGCAACAATAACATCGCTACTGTAGAGAACACACTTCAAG
 AACCTTTGGACCTCAGGGCTATACATGGATAGCAATTACCTGGACACGCTGTCCCCGG
 GAAATTGCGGGGCTGCAAAACCTAGAGTACCTGAACGTGGAGTACAACGCTATCCAGCTCA
 TCCTCCGGGACTTCAATGCCATGCCAAACTGAGGATCCTCATCTCAACAAACACCTG
 CTGAGGTCCCTGCCTGTGGACGTGTCGCTGGGTCTCGCTCTAAACTCAGCCTGCACAA
 CAATTACTCATGTACCTCCGGTGGCAGGGGTGCTGGACCAAGTTAACCTCCATCATCAGA
 TAGACCTCCACGGAAACCCCTGGGAGTGTGATGAGCGACCTCAAGTGTGAGACGCCGG
 GAACGCTTGGGTTCCGAAGTGTGATGAGCGACCTCAAGTGTGAGACGCCGGTAACCTCTT
 TAGAAAGGATTTCATGCTCTCTCAAATGACGAGATCTGCCCTCAGCTGTACGCTAGGATCT
 CGCCACGTTAACCTCGCACAGTAAAACAGCACTGGGTTGGCGAGACCGGGACGCACTCC
 AACTCCTACCTAGACACCAGCAGGGGTGTCATCGGTGTTGGTCCCGGGACTGCTGCTGG
 GTTTGTCACCTCCGCCTCACGTGGGGCATGCTCGTTATCCTGAGGAACCGAAAGC
 GGTCCAAGAGACGGAGATGCAACTCCTCGCGTCCGAGAGTTAACCTACAGACAGTCTGT
 GACTCTCCTACTGGCACAATGGGCTTACAACGAGATGGGCCCACAGAGTGTATGACTG
 TGGCTCTACTCGCTCTCAGACTAAAGACCCCAACCCAAATAGGGAGGGCAGAGGGAAAGGCG
 ATACATCCTCCCCACCGCAGGCACCCGGGGCTGGAGGGGCGTGTACCCAAATCCCGCG
 CCATCAGCCTGGATGGCATAAGTAGATAAAACTGTGAGCTCGCACAACCGAAAGGGCCT
 GACCCCTTACTTAGCTCCCTCTGAAACAAAGAGCAGACTGTGGAGAGCTGGAGAGCGCA
 GCCAGCTCGCTTGTGAGAGGCCCTTGACAGAAAGCCCAGCACGACCCCTGCTGGAAG
 AACTGACAGTGCCTCGCCCTCGGCCCCGGGGCTGTGGGGTTGGATGCCCGGGTTCTATAC
 ATATATACATATCCACATCTATATAGAGAGATAGATATCTATTTCCTGTGGATTAG
 CCCCGTGATGGCTCCCTGTTGGCTACGCAGGGATGGCAGTTGCACGAAGGCATGAATGTAT
 TGAAATAAGTAACCTTGACTTCTGAC

FIGURE 34

MLLWILLLETSLCFAAGNVTGVCKEKICSCNEIEGDLHVDCEKKGFTSLQRFTAPTSQFYH
LFLHGNSLTRLFPNEFANFYNAVSLHMENNGLHEIVPG AFLGLQLVKRLHINNNKIKSFRKQ
TFLGLDDLEYLQADFNLLRIDPGAFQDNLKLEVILNDNLISTLPANVFQYVPITHLDLRG
NRLKTLPYEEVLEQIPGIAEILLEDPWDCTCDLLSLKEWLENIPKNALIGRVVCEAPTRLQ
GKDLNETTEQDLCPLKNRVDSSLPA APPAQEEFTFAPGPLPTPKTNGQEDHATPGSAPNGGT
IPGNWQIKIRPTAAITGSSRNKPLANSLPCPGGCSDHIPGSGLKMNCNNRNVSSLADLKP
KLSNVQELFLRDNKIHSIRKSHFVDYKNLILLLGNNNIATVENNTFKNLLDLRWLYMDSNY
LDTLSREKFAGLQNLEYLNVEYNAIQLILPGTFNAMPKLRILILNNNNLLRSILPVDVFAGVSL
SKLSLHN NYFMYLPVAGVLDQLTSIIQIDLHGNPWECSCTIVPFKQWAERLGSEVLMSDLKC
ETPVNFRRKDFMLLSNDEICPQLYARISPTLTSHSKN STGLAETGTHSNSYLDTSRVSISVL
VPGLLL VFTSAFTVVGMLV FILRNRKRSKRRDANSSASEINSLQTVCDSSYWHNGPYNADG
AHRVYDCGSHSLSD

Signal sequence:

amino acids 1-15

Transmembrane domain:

amino acids 618-638

N-glycosylation site.

amino acids 18-22, 253-257, 363-367, 416-420, 595-599, 655-659

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 122-126, 646-650

Casein kinase II phosphorylation site.

amino acids 30-34, 180-184, 222-226, 256-260, 366-370, 573-577,
608-612, 657-661, 666-670, 693-697

N-myristoylation site.

amino acids 17-23, 67-73, 100-106, 302-308, 328-334, 343-349,
354-360, 465-471, 493-499, 598-604, 603-609

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 337-348

FIGURE 35

AGTCGACTGCGTCCCTGTACCCGGGCCAGCTGTGTTCTGACCCCAGAATAACTCAGGGC
 TGCACCGGGCCTGGCAGCGCTCCGACACATTCTGTGCGGCCCTAAGGGAAACTGTTGGC
 CGCTGGGCCGGGGGGATTCTTGGCAGTTGGGGGTCGCTGGGAGCGAGGGCGGAGGG
 AAGGGAGGGGAACCGGGTTGGGAAGCCAGCTGTAGAGGGCGGTGACCGCGCTCCAGACAC
 AGCTCTGCGTCCCTCGAGCGGGACAGATCCAAGTTGGGAGCAGCTCTGCGTGC
 CGGGAGAATGAGGCCGGCGTTGCCCTGTGCCCTCTGGCAGGCGCTCTGCCCGGGCGG
 CGCGAACACCCCCACTGCCGACCGTGTGGCTGCCCTCGGGGCTGCTACAGCCTGC
 ACCACGCTACCATGAAGCGGCAGGCCGGCAGGGAGGCTGCATCCTGCGAGGTGGGCGCTC
 AGCACCGTGCCTGGGGCGAGCTGCGCGTGTGCTCGCCTCTGCGGCCAGGCCAGG
 GCCCGAGGGGCTCAAAGACCTGCTGTTCTGGGTCGCACTGGAGCGCAGGCGTTCCACT
 GCACCTGGAGAACGAGCCTTGCGGGGTTCTCTGGCTGTCCCTCGACCCGGCGTCTC
 GAAAGCGACACGCTGCAGTGGTGGAGGAGCCCCAACGCTCCTGCACCGCGGAGATGCGC
 GGTACTCCAGGCCACCGTGGGTCGAGCCCGAGGCTGGAAGGAGATGCGATGCCACCTGC
 GCGCAACGGCTACCTGTCAAGTACCAAGTTGAGGTCTTGTGTCCTGCGCCGCCCCGG
 GCCGCTCTAACCTGAGCTATCGCGCCCTTCAGCTGACAGCGCCGCTCTGGACTTCAG
 TCCACCTGGGACCGAGGTGAGTGCCTCTGCCGGGACAGCTCCGATCTCAGTTACTGCA
 TCGCGACGAAATCGCGCTCGCTGGACAAACTCTCGGGGATGTGTTGTGTCCTGCCCC
 GGGAGGTACCTCCGTGGCAAATGCGCAGAGCTCCCTAACTGCCCTAGACGACTTGGGAGG
 CTTTGCCTGCGAATGTGCTACGGGCTTCAGCTGGGAAGGACGCCGCTCTGTGACCA
 GTGGGGAAAGGACAGCCGACCCCTGGGGGACCGGGGTGCCACCAAGCGCCGCCACT
 GCAACCAAGCCCCGTGCCAGAGAACATGGCCAATCAGGGTCGACGAGAACGACTGGGAG
 ACCACTTGTCCCTGAACAAGACAATTCACTGAGCTATCTATTGAGATTCTCGATGGGAT
 CACAGAGCACGATGTCTACCCCTCAAATGTCCCTCAAGCCGAGTCAAAGGCCACTATCACC
 CCATCAGGGAGCGTGATTCCAAGTTAATTCTACGACTCCTCTGCCACTCCTCAGGCTT
 CGACTCCTCTGCCGTGGTCTCATATTGTGAGCACAGCAGTAGTAGTGTGATCTGCGGG
 TGACCATGACAGTACTGGGCTTGCAAGCTCTGCTTACGAAAGCCCCCTTCCAGCCA
 AGGAAGGAGTCTATGGGCCGCCGGCTGGAGAGTGATCTGAGGCCGCTGCTTGGGCTC
 CAGTTCTGCACATTGACAAACAATGGGGTAAAGTCGGGACTGTGATCTGCGGG
 CAGAGGGTGCCTTGTGGCGAGTCCCTCTGGCTCTAGTGATGCA **TAGGAAACAGGG**
 CATGGGCACTCCTGTGAACAGTTTCACTTTGATGAAACGGGAACCAAGAGGAACCTAC
 TTGTGTAAGTACAATTCTGCAGAAATCCCCCTCTAAATCCCTTACTCCACTGAG
 GAGCTAAATCAGAACTGCACACTCCTCCCTGATGATGAGAGGAAAGTGGAAAGTGCCTT
 TGGTGTGAACTGGGGACCGGGTAGTGCTGGGAGAGATATTCTTATGTTATTGGAGAA
 TTTGGAGAAGTGATTGAACTTTCAAGACATTGAAACAAATAGAACACAATATAATTACA
 TTAAAAAAATAATTCTACCAAAATGGAAAGGAAATGTTCTATGTTGTCAGGCTAGGAGTAT
 ATTGGTTGAAATCCCAGGGAAAAAATAAAAATTAAGGATTGTTGAT

FIGURE 36

MRPAFALCLLWQALWPGPGGGEHPTADAGCSASGACYSLHHATMKRQAAEACILRGGALS
TVRAGAELRAVLALLRAGPGPGGGSKDLLFWVALERRSHCTLENEPLRGFSWLSSDPGGLE
SDTLQWVEEPQRSTCARRCAVLQATGGVEPAGWKEMRCHLRANGYLCKYQFEVLCPAPRPGA
ASNLSYRAPFQLHSAALDFSPPGTEVSALCRGQLPISVTCIADEIGARWDKLSGDVLCPCPG
RYLRAGKCAELPNCLDDLGGFACECATGFELGKDGRSCVTSSEGQPTLGGTGVPTRRPPATA
TSPVPQRTWPIRVDEKLGETPLVPEQDNSVTSIPEIPRWGSQSTMSTLQMSLQAESKATITP
SGSVISKFNSTTSSATPQAFDSSSAVVFIFVSTAVVVLVILTMVLGLVKLCFHESPSSQPR
KESMGPPGLESDPPEPAALGSSSAHCTNNGVKVGDCDLRDRAEGALLAESPLGSSDA

Signal sequence:

amino acids 1-16

Transmembrane domain:

amino acids 399-418

N-glycosylation site.

amino acids 189-193, 381-385

Glycosaminoglycan attachment site.

amino acids 289-293

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 98-102, 434-438

Casein kinase II phosphorylation site.

amino acids 275-279, 288-292, 342-346, 445-449

N-myristoylation site.

amino acids 30-36, 35-41, 58-64, 59-65, 121-127, 151-157,
185-191, 209-215, 267-273, 350-356, 374-380, 453-459, 463-469,
477-483

Aspartic acid and asparagine hydroxylation site.

amino acids 262-274

FIGURE 37

CGGACCGCGTGGGATTCAAGCAGTGGCCTGTGGCTGCCAGAGCAGCTCCTCAGGGGAAACTAAG
 CGTCGAGTCAGACGGCACCATATACTGCCCTTAAAGTGCCTCCGCCCTGCCGCCGTATC
 CCCCCGGCTACCTGGGCCGCCCCCGCGCGGTGCGCGGTGAGAGGGAGCGCGCGGGCAGCGA
 GCGCCGGTGTGAGCCAGCGCTGCTGCAGTGTGAGCGGGGTGTGAGCGCGGTGGGTGCGGA
 GGGGCGTGTGTGCCGGCGCGCGCCGTGGGTGCAAACCCCGAGCGTCTACGCTGCC**ATGA**
 GGGGCGCGAACGCCTGGCGCCACTCTGCCTGCTGCGCCACCCAGCTCTGCCGG
 CAGCAGTCCCCAGAGAGACCTGTTTACATGTGGTGGCATTCTTACTGGAGAGTCTGGATT
 TATTGGCAGTGAAGGTTTCTGGAGTGTACCTCCAAATAGCAAATGTACTGGAAAATCA
 CAGTTCCCAGAGAAAAGTAGTCGTTCTCAATTCCGATTCAAGACCTCGAGAGTGACAAC
 CTGTGCCGCTATGACTTGTGGATGTGTACAATGCCATGCCATGGCAGCGCATTGCCG
 CTTCTGTGGCACTTCCGGCTGGAGCCCTGTCAGTGGCAACAAGATGATGGTGCAGA
 TGATTCTGATGCCAACACAGCTGCCAATGGCTTCATGCCATGTTCTCGCTGCTGAACCA
 AACGAAAGAGGGATCAGTATTGTGGAGGACTCTTGACAGACCTCCGGCTTTAAAAC
 CCCCACGGCCAGACGGGATTACCGTGCAGGAGTCACTTGTGTGGCACATTGAGCCC
 CAAAGAACATCAGCTTATAGAATTAAAGTTGAGAAGTTGATGTGGAGCGAGATAACTACTGC
 CGATATGATTATGTGGCTGTGTTAATGGCGGGAACTCAACGATGCTAGAAGAAATTGAAA
 GTATTGTGGTGTAGTCCACCTGCGCCAATTGTCAGAGAGAAATGAACTTCTTATTCACT
 TTTTATCAGACTTAAGTTAACGCAAGATGGTTATTGGTCACTACATATTCAAGGCAAAA
 AAACGCCCTACAACACTACAGAACAGCCTGTCACCAACATTCCCTGTAACCACGGGTTAAA
 ACCCACCGTGGCCTTGTGTCACAAAAGTGTAGACGGACGGGACTCTGGAGGGCAATTATT
 GTTCAAGTGAATTGTATTAGCCGGACTGTTATCACAACCATACTCGCGATGGGAGTTG
 CACGCCACAGTCTCGATCATCACATCACAAAGAGGGAAATTGGCGATTCAAGCAGGGGG
 CAAGAACATGAGTGCAGGCTGACTGTCGTGCAAGCAGTGCCTCTCCTCAGAACAGGTC
 TAAATTACATTATTATGGGCCAAGTAGGTGAAGATGGCGAGGAAAATCATGCCAACAGC
 TTTATCATGATGTTCAAGACCAAGAACATCAGAACAGCTCTGGATGCCCTAAAAAATAGCAATG
TTAACAGTGAACGTGTCATTAAAGCTGTATTCTGCCATTGCCCTTGAAAGATCTATGTT
 TCTCAGTAGAAAAAAATACTTATAAAATTACATATTGAAAGAGGATCCGAAAGATGG
 GACTGGTTGACTCTCACATGATGGAGGTATGAGGCCTCCGAGATAGCTGAGGGAAAGTCTT
 TGCCTGCTGTCAGAGGAGCAGCTATCTGATTGGAAACCTGCCACTAGTGCAGGTGATAGGA
 AGCTAAAAGTGTCAAGCGTTGACAGCTGGAAGCGTTATTATACATCTGTAAAAGGAT
 ATTTTAGAATTGAGTTGTGTGAAGATGTCAAAAAAAGATTTAGAAGTGCACATTTATAGT
 GTTATTGTTTACCTTCAAGCCTTGCCCTGAGGTGTTACAATCTGTCTTGCCTTCTA
 AATCAATGCTTAATAAAATATTTAAAGGAAAAAA

FIGURE 38

MRGANAWAPLCLLLAAATQLSRQQSPERPVFTCGGILTGESGFIGSEGFPGVYPPNSKCTWK
ITVPEGKVVVLNFRFIDLESDNLCRYDFVDVYNGHANGQRIGRFCGTFRPGALVSSGNKMMV
QMISDANTAGNGFMAMFSAAEPNERGDQYCGGLLDRPSGSFKTPNWPDRDYPAGVTCVWHIV
APKNQLIELKFEKFDVERDNYCRYDYVAVFNGEVNDARRIGKYCGDSPAPIVSERNELLI
QFLSDLSLTADGFIGHYIFRPKKLPTTTEQPVTTFPVTTGLKPTVALCQQKCRRTGTLEGN
YCSSDFVLAGTVITTRDGLHATVSIINIYKEGNLAIQQAGKNMSARLTVVCKQCPLLRR
GLNYIIMGQVGEDGRGKIMPNSFIMMFKTKNQKLLDALKNKQC

Signal sequence:

amino acids 1-23

N-glycosylation site.

amino acids 355-359

Casein kinase II phosphorylation site.

amino acids 64-68, 142-146, 274-278

Tyrosine kinase phosphorylation site.

amino acids 199-208

N-myristoylation site.

amino acids 34-40, 35-41, 100-106, 113-119, 218-224, 289-295,
305-311, 309-315, 320-326, 330-336

Cell attachment sequence.

amino acids 149-152

FIGURE 39

CGGACGCGTGGCGGACGCGTGGCGGCCACGGCGCCCGGGCTGGGCGGTCGCTCTT
CCTTCTCCGTGGCTACGAGGGTCCCCAGCCTGGTAAAGATGGCCCCATGGCCCCGAAGG
GCCTAGTCCCAGCTGTGCTCTGGGCCTCAGCCTCTCCTCAACCTCCCAGGACCTATCTGG
CTCCAGCCCTCTCCACCTCCCCAGTCTCTCCCCGCCTCAGCCCCATCCGTGTACACCTG
CCGGGGACTGGTTGACAGCTTAACAAGGGCTGGAGAGAACCATCCGGACAACTTGGAG
GTGGAAACACTGCCTGGGAGGAAGAGAATTTGCCAAATACAAAGACAGTGAGACCCGCCTG
GTAGAGGTGCTGGAGGGTGTGCAGCAAGTCAGACTCGAGTGCCACCGCCTGCTGGAGCT
GAGTGAGGAGCTGGTGGAGAGCTGGTGGTTACAAGCAGCAGGAGGCCGGACCTCTTCC
AGTGGCTGTGCTCAGATTCCCTGAAGCTCTGCTGCCCGCAGGCACCTCGGCCCTCCTGC
CTTCCCTGTCTGGGAAACAGAGAGGCCCTGGCTACGGGAGTGTGAAGGAGAAGG
GACACGAGGGGGCAGCGGGCACTGTGACTGCCAAGCCGGCTACGGGGTGAGGCCTGTGGCC
AGTGTGGCCTTGGCTACTTGAGGCAGAACGCAACGCCAGCCATCTGGTATGTTCGGCTTGT
TTTGGCCCTGTGCCGATGCTCAGGACCTGAGGAATCAAACGTGTTGCAATGCAAGAAGGG
CTGGGCCCTGCATCACCTCAAGTGTAGACATTGATGAGTGTGGCACAGAGGGAGCCA
GTGGAGCTGACCAATTCTCGTGAACACTGAGGGCTCCTATGAGTGCCGAGACTGTGCCAAG
GCCTGCCTAGGCTGCATGGGGCAGGGCAGGTGCTGTAAGAAGTGTAGCCCTGGCTATCA
GCAGGTGGCTCCAAGTGTCTCGATGTGGATGAGTGTGAGACAGAGGTGTCCGGAGAGA
ACAAGCAGTGTGAAAACACCGAGGGCGTTATCGCTGCATCTGTGCCGAGGGCTACAAGCAG
ATGGAAGGCATCTGTGTGAAGGAGCAGATCCCAGAGTCAGCAGGCTTCTCAGAGATGAC
AGAAGACGAGTTGGTGGTGCAGCAGATGTTGGCATCATCTGTGCACTGGCCA
CGCTGGCTGCTAAGGGCGACTTGGTGGTACCGCCATCTCATTGGGCTGTGGCGGCCATG
ACTGGCTACTGGTGTCAAGAGCGCAGTGACCGTGTGGAGGGCTTCATCAAGGGCAGATA
ATCGCGGCCACCAACCTGTAGGACCTCTCCACCCACGCTGCCAGAGCTTGGCTGCC
TCCTGCTGGACACTCAGGACAGCTTGGTTATTTGAGAGTGGGTAAGCACCCCTACCTG
CCTTACAGAGCAGCCCAGGTACCCAGGCCGGCAGACAAGGCCCTGGGTAAAAAGTAGC
CCTGAAGGTGGATACCATGAGCTCTCACCTGGGGACTGGCAGGCTTCACAATGTGTGA
ATTCAAAAGTTTCTTAATGGTGGCTGCTAGAGCTTGGCCCTGCTTAGGATTAGGTG
GTCCTCACAGGGTGGGCCATCACAGCTCCCTCTGCCAGCTGCATGCCAGTTCTGT
TCTGTGTTCACCAACATCCCCACACCCATTGCCACTTATTATTCATCTCAGGAAATAAGA
AAGGTCTTGGAAAGTTAAAAAAAAAAAAAAAAAAAAAA

FIGURE 40

MAPWPPKGLVPAVLWGLSLFLNLPGPIWLQPSPPPQSSPPQPHPCHTCRGLVDSFNKGLER
TIRDNFGGGNTAWEEEENLSKYKDSETRLVEVLEGVCSKSDFECHRLLELSEELVESWWFHKGQ
QEAPDLFQWLCSDSLKLCCPAGTFGPSCLPCCPGGTERPCGGYGQCEGEGTRGGSGHCDCQAG
YGGECACGQCGLGYFEAERNASHLVCACFGPCARCSGPEESNCLQCKKGWALHHLKCVQIDE
CGTEGANCGADQFCVNTEGGSYECRDCAKACLGMGAGPGRKKCSPGYQQVGSKCLDVDECE
TEVCPGENKQCENTEGGYRCICAEGYKQMEGICVKEQI PESAGFFSEMTEDELVVLQQMFFG
IIICALATLAAKGDLVFTAIFIGAVAAMTGYWLSERSDRVLEGFIKGR

Signal sequence:

amino acids 1-29

Transmembrane domain:

amino acids 372-395

N-glycosylation site.

amino acids 79-83, 205-209

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 290-294

Casein kinase II phosphorylation site.

amino acids 63-67, 73-77, 99-103, 101-105, 222-226, 359-263

N-myristoylation site.

amino acids 8-14, 51-57, 59-65, 69-75, 70-76, 167-173, 173-179,
177-183, 188-194, 250-256, 253-259, 267-273, 280-286, 283-289,
326-332, 372-378, 395-401

Aspartic acid and asparagine hydroxylation site.

amino acids 321-333

EGF-like domain cysteine pattern signature.

amino acids 181-193

FIGURE 41

TGAGACCCTCCTGCAGCCTCTCAAGGGACAGCCCCACTCTGCCTCTGCTCCTCCAGGGCA
 GCACCATGCAGCCCCCTGTGGCTCTGCTGGGCACTCTGGGTGTTGCCCTGGCCAGCCCCGGG
 GCCGCCCTGACCGGGGAGCAGCTCCTGGCAGCCTGCTGCCAGCTGCAGCTCAAAGAGGT
 GCCCACCCCTGGACAGGGCCGACATGGAGGAGCTGGTCATCCCCACCCACGTGAGGGCCCAGT
 ACGTGGCCCTGCTGCAGCGCAGCCACGGGGACCCTCCCGGGAAAGAGGTTCAGCCAGAGC
 TTCCGAGAGGTGGCCGGCAGGTTCTGGCGTTGGAGGCCAGCACACACCTGCTGGTGGTCGG
 CATGGAGCAGCGGCTGCCGCCAACAGCGAGCTGGTCAGGCCGTGCTGCCGTCTTCAGG
 AGCCGGTCCCCAAGGCCCGCTGCACAGGCACGGGGGGCTGTCCCCCGCGAGCGCCCGGGCC
 CGGGTGACCGTCAGTGGCTGCGCGTCCCGACGACGGCTCCAACCGCACCTCCCTCATCGA
 CTCCAGGCTGGTGTCCGTCCACGAGAGCGGCTGGAAGGCCTTCGACGTGACCGAGGCCGTGA
 ACTTCTGGCAGCAGCTGAGCCGGCCCGGCAGCCGCTGCTGCTACAGGTGTCGGTGCAGAGG
 GAGCATCTGGGCCGCTGGCGTCCGGCCACAAGCTGGTCCGCTTGCCCTCGCAGGGGC
 GCCAGCCGGCTTGGGAGCCCCAGCTGGAGCTGCACACCCCTGGACCTTGGGACTATGGAG
 CTCAGGGCGACTGTGACCTGAAGCACCAATGACCGAGGGCACCCGCTGCTGCCGCCAGGAG
 ATGTACATTGACCTGCAGGGATGAAGTGGCCGAGAACTGGGTGCTGGAGCCCCCGGGCTT
 CCTGGCTTATGAGTGTGTCGGCACCTGCCGGAGCCCCCGGAGGCCCTGGCCTCAAGTGGC
 CGTTCTGGGCCCTGACAGTGCATGCCCTGGAGACTGACTCGCTGCCATGATGTCAGC
 ATCAAGGAGGGAGGCAGGACCAGGCCAGGTGGTCAGCCTGCCAACATGAGGGTGCAGAA
 GTGCAGCTGTGCCCTGGATGGTGCCTCGTGCAGGAGGCTCCAGCCATGGCCCTAGTG
 TAGCCATCGAGGGACTTGACTTGTGTGTTCTGAAGTGTGAGGGTACCAAGGAGAGCTG
 GCGATGACTGAAGTGCATGGACAAATGCTCTGTGCTCTAGTGAGCCCTGAATTGCTT
 CCTCTGACAAGTTACCTCACCTAATTTGCTCTCAGGAATGAGAATCTTGGCCACTGGA
 GAGCCCTGCTCAGTTCTTCTATTCTTATTCACTGCACTATATTCAAGCACTTACAT
 GTGGAGATACTGTAACCTGAGGGCAGAAAGCCANTGTGTCATTGTTACTGTCTGTCAC
 TGGATCTGGCTAAAGTCCACCACCTGGACCTAACAGACCTGGGTTAAGTGTGGGT
 TGTGCATCCCCAATCCAGATAATAAGACTTGTAAAACATGAATAAACACATTATTCT
 AAAA

FIGURE 42

MQPLWLCWALWVLPLASPGAAALTGEQLLGSLLRQLQLKEVPTLDRADMEELVIPHRAQYV
ALLQRSHGDRSRGKRFQSFSREVAGRFLALEASTHLLVFGMEQRLPPNSELVQAVLRLFQEP
VPKAALHRHGRLSPRSARARVTVEWLRVRDDGSNRTSLIDSRLVSVHESGWKAFDVTEAVNF
WQQLSRPRQPLLLQSVQREHLGPLASGAHKLVRFASQGAPAGLGEPOLELHTLDLGDYGAQ
GDGDPPEAPMTEGTRCCRQEMYIDLQGMKWAENWVLEPPGFLAYECVGTCRQPPEALAFKWPF
LGPRQCIASETDSLPMIVSIKEGGRTRPQVVSLPNMRVQKCSCASDGALVPRRLQP

Signal sequence:

amino acids 1-18

N-glycosylation site.

amino acids 158-162

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 76-80

Casein kinase II phosphorylation site.

amino acids 68-72, 81-85, 161-165, 169-173, 319-323, 329-333

N-myristoylation site.

amino acids 19-25, 156-162, 225-231, 260-266, 274-280

Amidation site.

amino acids 74-78

TGF-beta family signature.

amino acids 282-298

FIGURE 43

GTCTGTTCCCAGGAGTCCTCGCGGCTGTTGTCAGTGGCCTGATCGCGATGGGGACAAA
GGCGCAAGTCGAGAGGAAACTGTTGCGCTCTCATATTGGCGATCCTGTTGCTCCCTGG
CATTGGGCAGTGTACAGTGCACCTCTGAACCTGAAGTCAGAATTCTGAGAATAATCCT
GTGAAGTTGTCCTGCGCTACTCGGGCTTTCTCTCCCCGTGTGGAGTGGAGTTGACCA
AGGAGACACCACCAACTCGTTGCTATAATAACAAGATCACAGCTCCTATGAGGACCGGG
TGACCTTCTTGCCAACCTGGTATCACCTCAAGTCCGTACACGGGAAGACACTGGGACATAC
ACTTGTATGGTCTCTGAGGAAGGCGGAAACAGCTATGGGGAGGTCAAGGTCAAGCTCATCGT
GCTTGTGCCTCCATCCAAGCCTACAGTTAACATCCCCCTCCTCTGCCACCATTGGAACCGGG
CAGTGCTGACATGCTCAGAACAGATGGTCCCCACCTTCTGAATAACACCTGGTCAAAGAT
GGGATAGTGTGATGCCCTACGAATCCAAAAGCACCCGTGCCCTCAGCAACTCTCCTATGTCCT
GAATCCCACACAGGAGAGCTGGTCTTGATCCCCTGTCAGCCTCTGATACTGGAGAATACA
GCTGTGAGGCACGGAATGGGTATGGGACACCCATGACTCAAATGCTGTGCGCATGGAAGCT
GTGGAGCGGAATGTGGGGTCATCGTGGCAGCCGTCTGTAACCTGATTCTCCTGGGAAT
CTTGGTTTTGGCATCTGGTTGCCTATAGCCGAGGCCACTTGACAGAACAAAGAAAGGGA
CTTCGAGTAAGAAGGTGATTACAGCCAGCCTAGTGCCCGAAGTGAAGGAGAATTCAAACAG
ACCTCGTCATTCCCTGGTGTAGCCTGGTCGGCTACCGCCTATCATCTGCATTGCTTACT
CAGGTGCTACCGGACTCTGCCCTGATGTCTGTAGTTCACAGGATGCCCTATTGCTTC
TACACCCACAGGGCCCCCTACTTCTCGGATGTGTTTAATAATGTCAGCTATGTGCC
ATCCTCCTTCATGCCCTCCCTCCCTTACCACTGCTGAGTGGCCTGAACTTGTAA
GTGTTATTCCCATTCTTGAGGGATCAGGAAGGAATCCTGGGTATGCCATTGACTTCCC
TTCTAAGTAGACAGAAAAATGGCGGGGTCCAGGAATCTGCACTCAACTGCCACCTGGC
TGGCAGGGATCTTGAATAGGTATCTTGAGCTGGTCTGGCTCTTCTGTACTGAC
GACCAGGGCCAGCTGTTCTAGAGCGGAATTAGAGGCTAGAGCGGCTGAAATGGTTGG
TGATGACACTGGGTCTTCATCTCTGGGCCACTCTCTGTCTTCCATGGGAAGTG
CCACTGGGATCCCTCTGCCCTGCTCTGAATAACAGCTGACTGACATTGACTGTCTGT
GGAAAATGGAGCTTGTGGAGAGCATAGTAAATTTCAGAGAACCTGAAGCCAAAAG
GATTAAAACCGCTGCTCTAAAGAAAAGAAAATGGAGGCTGGCGCAGTGGCTACGCC
TAATCCCAGAGGCTGAGGCAGGCAGGATCACCTGAGGTGGAGTTGGGATCAGCCTGACCA
ACATGGAGAACCTACTGGAAATACAAAGTTAGCCAGGCATGGTGGTGCATGCCGTAGTC
CCAGCTGCTCAGGAGCCTGGCAACAAGAGCAAAACTCCAGCTCAAAAAAAAAAAAAAA

FIGURE 44

MGTKAQVERKLLCLFILAILLCSLALGSVTVHSSEPEVRIPENNPVKLSCAYSGFSSPRVEW
KFDQGDTTRLVCYNNKITASYEDRVTFLPTGITFKSVTREDTGTYTCMVSEEGGNSYGEVKV
KLIVLVPPSKPTVNIPSSATIGNRAVLTCSSEQDGSPPSEYTWFKDGIIVMPTNPKSTRAFSNS
SYVLNPTTGEVFDPLSASDTGEYSCEARNGYGTPMTSNAVRMEAVERNVGVIVAAVLVTLI
LLGILVFGIWFAYSRGHFDRTKKGTSSKKVIYSQPSARSEGEFKQTSSFLV

Signal sequence:

amino acids 1-27

Transmembrane domain:

amino acids 238-255

N-glycosylation site.

amino acids 185-189

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 270-274

Casein kinase II phosphorylation site.

amino acids 34-38, 82-86, 100-104, 118-122, 152-156, 154-158,
193-197, 203-207, 287-291

N-myristylation site.

amino acids 105-111, 116-122, 158-164, 219-225, 237-243, 256-262

FIGURE 45

CAGCGCGTGGCCGGCGCCGCTGTGGGACAGC**ATG**AGCGGCGGTTGGATGGCGCAGGTTGGA
GCGTGGCGAACAGGGCTCTGGCCTGGCGCTGCTGCTGCTCGGCCTCGACTAGGCCT
GGAGGCCGCGAGCCGCTTCACCCGACCTCTGCCAGGCCAGGCCAGCTCAG
GCTCGTCCCACCAAGTCCAGTGCCGACCAAGTGGCTTATCGTGCCCTCACCTGG
CGCTGCGACAGGGACTTGGACTGCAGCGATGGCAGCGATGAGGAGGTGCAGGATTGAGCC
ATGTACCCAGAAAGGGCAATGCCACGCCCTGGCTCCCTGCCCTGCACCGCGTCA
GTGACTGCTCTGGGGAACTGACAAGAAACTGCGCAACTGCAGCCCTGGCTGCCTAGCA
GGCGAGCTCGTTGCACGCTGAGCGATGACTGCATTCCACTCACGTGGCGCTGCACGGCCA
CCCAGACTGTCCCAGCTCCAGCGACGAGCTGGCTGTGGAACCAATGAGATCCTCCGGAAAG
GGGATGCCACAACCAGGGGCCCCCTGTGACCCCTGGAGAGTGTACCTCTCAGGAATGCC
ACAACCAGGGGCCCCCTGTGACCCCTGGAGAGTGTCCCTCTGTGGGAATGCCACATCCTC
CTCTGCCGGAGACCAGTCTGGAAGCCAAC TGCCCTATGGGTTATTGCAGCTGCTGCCGTGC
TCAGTGCAAGCCTGGTCACCGCCACCCCTCCTCTTGTCCCTGGCTCCGAGCCAGGAGCGC
CTCCGCCACTGGGTTACTGGTGGCCATGAAGGAGTCCCTGCTGTCAGAACAGAAC
CTCGCTGCC**TG**AGGACAAGCAACTGCCACCCACCGTCACTCAGCCCTGGCGTAGCCGGACA
GGAGGAGAGCAGTGATGCGATGGGTACCCGGCACACCAGCCCTCAGAGACCTGAGTTCTT
CTGGCCACGTGGAACCTCGAACCCGAGCTCCTGCAGAAGTGGCCCTGGAGATTGAGGGTCCC
TGGACACTCCCTATGGAGATCCGGGAGCTAGGATGGGAACCTGCCACAGCCAGAAC
GGGCTGGCCCCAGGCAGCTCCAGGGGTAGAACGCCCTGTGCTTAAGACACTCCCTGCTG
CCCCGTCTGAGGGTGGCGATTAAGTTGCTTC

FIGURE 46

MSGGWMAQVGAWRTGALGLALLLGLGLEAAASPLSTPTSAQAAGPSSGSCPPTKFQCR
TSGLCVPLTWRCRDLDSDGSDEEECRIEPCTQKGQC PPPGLPCPCTGVSDCSGGTDKKL
RNCsRLACLAGELRCTLSDDCIPLTWRCDGHPDCPDSSDELGCGTNEILPEGDATTMGPPVT
LESVTSLRNATTMGPPTLESVPSVGNATSSAGDQSGSPTAYGVIAAAAVLSASLVTATLL
LLSWLRAQERLRPLGLLVAMKESLLLSEQKTSLP

Signal sequence:

amino acids 1-30

Transmembrane domain:

amino acids 230-246

N-glycosylation site.

amino acids 126-130, 195-199, 213-217

Casein kinase II phosphorylation site.

amino acids 84-88, 140-144, 161-165, 218-222

N-myristoylation site.

amino acids 3-9, 10-16, 26-32, 30-36, 112-118, 166-172, 212-218,
224-230, 230-236, 263-269

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 44-55

Leucine zipper pattern.

amino acids 17-39

FIGURE 47

CCCACCGTCCGGCTCGCTCGCGCAGCGGCCAGCAGAGGTCGCGCACAGATGCGG
GTTAGACTGGCGGGGGAGGAGGCGGAGGAGGAAGGAAGCTGCATGCATGAGACCCACAGA
CTCTTGAAGCTGGATGCCCTCTGTGGATGAAAGATGTATCATGGAATGAACCCGAGCAATG
GAGATGGATTCTAGAGCAGCAGCAGCAGCAGCAACCTCAGTCCCCCAGAGACTCTG
GCCGTGATCCTGTGGTTCAGCTGGCCTGTGCTTCGGCCCTGCACAGCTCACGGCGGGTT
CGATGACCTCAAGTGTGCTGACCCGGCATTCCGAGAATGGCTTCAGGACCCCCAGCG
GAGGGGTTTCTTGAAGGCTCTGTAGCCGATTTCACTGCCAAGACGGATTCAAGCTGAAG
GGCGCTACAAAGAGACTGTGTTGAAGCATTAAATGGAACCCTAGGCTGGATCCCAAGTGA
TAATTCCATCTGTGCAAGAAGATTGCCGTATCCCTCAAATCGAAGATGCTGAGATTATA
ACAAGACATATAGACATGGAGAGAACGTAATCATCACTTGTATGAAGGATTCAAGATCCG
TACCCGACCTACACAATATGGTTTCAATTATGTCGCGATGGAACGTGGAATAATCTGCC
CATCTGTCAAGGCTGCCTGAGACCTCTAGCCTCTTAATGGCTATGTAACACATCTCTGAGC
TCCAGACCTCCTTCCGGTGGGACTGTGATCTCCTATCGCTGCTTCCGGATTAAACTT
GATGGGTCTGCGTATCTTGAGTGCTTACAAAACCTTATCTGGTCGTCAGCCCACCCGGTG
CCTTGCTCTGGAAGCCCAAGTCTGTCACCTCAATGGTAGTCACGGAGATTCGTCT
GCCACCCGGCCTTGTGAGCGTACAACCACGGAACGTGGTGGAGTTTACTGCGATCCT
GGCTACAGCCTCACCAGCAGTACAAGTACATCACCTGCCAGTATGGAGAGTGGTTCCCTC
TTATCAAGTCTACTGCATCAAATCAGAGCAAACGTGGCCAGCACCCATGAGACCCCTCTGA
CCACGTGGAAGATTGTGGCGTTCACGGCAACCAGTGTGCTGCTGGTGCTGCTCGTCATC
CTGGCCAGGATGTTCCAGACCAAGTTCAAGGCCACTTCCCCCAGGGGCGTCCCCGGAG
TTCCAGCAGTGACCCCTGACTTGTGGTAGACGGCGTCCCCGTATGCTCCGTCTATG
ACGAAGCTGTGAGTGGCGGTTGAGTGCCTTAGGCCCGGTACATGGCCTCTGTGGCCAG
GGCTGCCCTTACCCGTGGACGACCAGAGCCCCCAGCATACCCGGCTCAGGGACACGGA
CACAGGCCAGGGAGTCAGAAACCTGTGACAGCGTCTCAGGCTTCTGAGCTGCTCCAAA
GTCTGTATTCACCTCCAGGTGCCAAGAGAGCACCCACCCCTGCTCGGACAACCTGACATA
ATTGCCAGCACGGCAGAGGAGGTGGCATCCACCAGCCAGGCATCCATCATGCCACTGGGT
GTTGTTCTAAGAAACTGATTGATTAAGGAAACTTCCAAAGTGTCCCTGAAGTGTCTCTCAA
ATACATGTTGATCTGTGGAGTTGATTCTTCTCTTGTGGTTAGACAAATGTAACAA
AGCTCTGATCCTAAAATTGCTATGCTGATAGAGTGGTAGGGCTGGAAGCTGATCAAGTC
CTGTTCTTCTTGACACAGACTGATTAAGGAAACTTAAAGNAAGNAAGNAAGNAAGNAAG

FIGURE 48

MYHGMNPSNGDGFLEQQQQQQQQPQSPQRLLAVILWFQLALCFGPAQLTGGFDDLQVCADPGI
PENGFRTPSGGVFFEGSVARFHQCQDGFKLKKGATKRLCLKHFNGTLGWI PSDNSICVQEDCRI
PQIEDAEIHNKTYRHGEKLIITCHEGFKIRYPDLHNMVSLCRDDGTWNNLPIQGCLRPLAS
SNGYVNISELQTSFPVGTVISYRCFPGFKLDGSAYLECLQNLIWSSSPRCLALEAQVCPLP
PMVSHGDFVCHPRPCERYNHGTVEFYCDPGYSLTSDYKYITCQYGEWFPSYQVYCIKSEQT
WPSTHETLLTTWKIVAFTATSVLLVLLVILARMFQTKFKAHFPPRGPPRSSSDPDFVVVD
GVPVMLPSYDEAVSGGLSALGPGYMASVGQGCPLPVDDQSPPAYPGSGDTDTGPGESETCDS
VSGSSELLQSLYSPPRCQESTHPASDNPDIIASTAAEVASTSPGIHHAHWVLFLRN

Signal sequence:

amino acids 1-41

Transmembrane domain:

amino acids 325-344

N-glycosylation site.

amino acids 104-108, 134-138, 192-196

Casein kinase II phosphorylation site.amino acids 8-12, 146-150, 252-256, 270-274, 313-317, 362-366,
364-368, 380-384, 467-471, 468-472**N-myristoylation site.**amino acids 4-10, 61-67, 169-175, 203-209, 387-393, 418-424,
478-484**Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 394-405

FIGURE 49

CCCACGCGTCCGCTCCGCGCCCTCCCCCCCACCGCTCCGTGCGGTCCGTCGGCTAGAGA
TGCTGCTGCCGCCGGTTGCAGTTGTCGCGCACGCCCTGCCGCCAGCCGCTCCACCGCCGT
AGCGCCCGAGTGTGGGGGGCGCACCCGAGTCGGGCCATGAGGCCGGAACCGCGCTACAGG
CCGTGCTGCCGTGCTGGGTGGCTGCGGCGACGGGCGCTGAGTGCCTGCTGAGTGCC
TCGGATTGGACCTCAGAGGAGGGCAGCCAGTCTGCCGGGAGGGACACAGAGGCCTTGTAA
TAAAGTCATTTACTTCATGATACTTCGAAGACTGAACACTTGAGGAAGCCAAAGAACGCT
GCAGGAGGGATGGAGGCCAGCTAGTCAGCATCGAGTCTGAAGATGAACAGAAACTGATAGAA
AAGTTCATTGAAAACCTCTGCCATCTGATGGTGAATTCTGGATTGGCTCAGGAGGCCTGA
GGAGAAACAAAGCAATAGCACGCCGCCAGGACCTTATGCTGGACTGATGGCAGCATAT
CACAATTAGGAACGGTATGTGGATGAGCCGCTGCCAGCGAGGTCTGCGTGGTCA
TACCATCAGCCATCGGCACCCGCTGGCATCGGAGGCCCTACATGTTCCAGTGAATGATGA
CCGGTGCAACATGAAGAACAAATTGCAAATATTCTGATGAGAACAGCAGTTCTT
CTAGAGAAGCTGAAGGTGAGGAAACAGAGCTGACAACACCTGTACTTCCAGAAGAACACAG
GAAGAAGATGCCAAAAAAACATTAAAGAAAGTAGAGAACGCTGCCTGAATCTGGCTACAT
CCTAATCCCCAGCATTCCCCCTCTCCCTCCTGTGGTCACCACAGTTGTATGTTGGTT
GGATCTGTAGAAAAAGAAAACGGGAGCAGCCAGACCTAGCACAAAGAACACACCAC
TGGCCCTCTCCTCACCAGGGAAACAGCCGGACCTAGAGGTCTACAATGTCATAAGAAAACA
AAGCGAAGCTGACTTAGCTGAGACCCGGCCAGACCTGAAGAATATTCTCGAGTGTGTT
CGGGAGAAGCCACTCCGATGACATGTCTGTGACTATGACAACATGGCTGTGAACCCATCA
GAAAGTGGTTGTGACTCTGGTGGAGAGTGGATTGTGACCAATGACATTATGA
GTTCTCCCCAGACCAAATGGGAGGACTAAGGAGTCTGGATGGTGAAAATGAAATATATG
GTTATTAGGACATATAAAACTGAAACTGACAACAATGAAAAGAAATGATAAGCAAATC
CTCTTATTTCTATAAGGAAAATACACAGAACGGTCTATGACAAGCTTAGATCAGGTCTGT
GGATGAGCATGTGGTCCCCACGACCTCTGTGGACCCCCACGTTGGCTGTATCCTTAT
CCCAGCCAGTCAGCTGACCTATGAGAACGGTACCTTGCCAGGTCTGGCACATAGTA
GAGTCTCAATAATGTCACTTGGTTGGTTGTATCTAACTTTAACGGACAGAGCTTACCTG
GCAGTGATAAAGATGGGCTGTGGAGCTGGAAAACCACCTCTGTTTCCCTGCTCTACAG
CAGCACATATTATCATAACAGACAGAAAATCCAGAACATCTTCAAAGCCCACATATGGTAGCACAG
GTTGGCCTGTGCATGGCAATTCTCATATCTGTTTTCAAAGAATAAAATCAAATAAAGA
GCAGGAAAAAA

FIGURE 50

MRPGTALQAVLLAVLLVGLRAATGRLLSASDLRGGQPVCRGGTQRPCYKVIYFHDT\$RRL
NFEEAKEACRRDGGQLVSIESEDEQKLIENLLPSDGF\$WIGLRRREEKQSNSTACQDL
YAWTDGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKY
SDEKPAVPSREAEGEETELTPVLPETQEEAKKTFKESREAALNAYILIPSIPLLLL
VTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSEADLAETRPDL
KNISFRVCSGEATPDDMSCDYDNMAVNPSSEGFVTLVSVESGFVTNDIYEFS\$PDQMGRSKES
GWVENEIYGY

Signal sequence:

amino acids 1-21

Transmembrane domain:

amino acids 235-254

N-glycosylation site.

amino acids 117-121, 312-316

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 296-300

Casein kinase II phosphorylation site.

amino acids 28-32, 30-34, 83-87, 100-104, 214-218, 222-226,
299-303, 306-310, 323-327

N-myristylation site.

amino acids 18-24, 37-43, 76-82, 146-152

FIGURE 51

GGGGTCTCCCTCAGGGCCGGGAGGCACAGCGGTCCCTGCTTGTGAAGGGCTGGATGTACGC
ATCCGCAGGTTCCCGCGGACTTGGGGCGCCCGCTGAGCCCCGGCGCCCGCAGAAGACTTGT
GTTTGCCTCCTGCAGCCTCAACCCGGAGGGCAGCGAGGGCCTACCACCATGATCACTGGTGT
GTTCAGCATGCGCTTGTGGACCCCAGTGGCGTCTGACCTCGCTGGCGTACTGCCTGCACC
AGCGGCGGGTGGCCCTGGCCGAGCTGCAGGAGGCCATGCCAGTGTCCGGTCACCGCAGC
CTGCTGAAGTTGAAAATGGTCAGGTGTTGACACGGGGCTGGAGTCCCTCAAGCC
GCTCCCGCTGGAGGGAGCAGGTAGAGTGAACCCCCAGCTATTAGAGGTCCCACCCAAACTC
AGTTGATTACACAGTCACCAATCTAGCTGGTGGTCCGAAACCATAATTCTCCTTACGACTCT
CAATACCACATGAGACCACCCCTGAAGGGGGCATGTTGCTGGCAGCTGACCAAGGTGGCAT
GCAGCAAATGTTGCCTTGGGAGAGAGACTGAGGAAGAACTATGTGGAAGACATTCCCTTC
TTTCACCAACCTCAACCCACAGGAGGTCTTATTGTTCCACTAACATTTCGGAATCTG
GAGTCCACCCGTTGTTGCTGGCTGGCTTTCCAGTGTCAAGAAAGAAGGCCATCATCAT
CCACACTGATGAAGCAGATTCAAGTCTGTATCCAACTACCAAAGCTGCTGGAGCCTGA
GGCAGAGAACAGAGGCCGAGGCAGACTGCCTCTTACAGCCAGGAATCTCAGAGGATTG
AAAAAGGTGAAGGACAGGATGGCATTGACAGTAGTGTAAAGTGGACTTCTCATCCTCCT
GGACAACGTGGCTGCCGAGCAGGCACACAACCTCCAAAGCTGCCCATGCTGAAGAGATTG
CACGGATGATCGAACAGAGAGACTGTGGACACATCCTGTACATACTGCCAAGGAAGACAGG
GAAAGTCTTCAGATGGCAGTAGGCCATTCCACATCCTAGAGAGCAACCTGCTGAAGC
CATGGACTCTGCCACTGCCCGACAAGATCAGAAAGCTGTATCTATGCCGCTCATGATG
TGACCTTCATACCGCTTTAATGACCTGGGATTTGACCACAAATGCCACCGTTGCT
GTTGACCTGACCATGGAACCTTACCAAGCACCTGGAATCTAAGGAGTGGTTGTGCAGCTCTA
TTACCACGGGAAGGAGCAGGTGCCGAGAGGTTGCCCTGATGGGCTCTGCCGCTGGACATGT
TCTTGAATGCCATGTCAGTTATACCTTAAGCCCAGAAAAATACCATGCACTCTGCTCTCAA
ACTCAGGTGATGGAAGTTGAAATGAAGAGTAACTGATTTATAAAAGCAGGATGTGTTGATT
TTAAAATAAGTGCCTTATACAATG

FIGURE 52

MITGVFSMRLWTPGVLTSLAYCLHQRRVALAELQEADGQCPVDRSLLKLKMVQVVFRHGAR
SPLKPLPLEEQVEWNPQLLEVPPQTQFDYTVTNLAGGPKPYSFYDSQYHETTLKGGMFAGQL
TKVGMQQMFALGERLRKNYVEDIPFLSPTFNPQEVFIRSTNIFRNLESTRCLLAGLFQCQKE
GPIIIHTDEADSEVLYPNYQSCWSLRQRTRGRRQTASLQPGISEDLKKVKDRMGIDSSDKVD
FFILLDNVAAEQAHNLPSCPMLKRFARMIEQRAVDTSLYILPKEDRESLQMAVGFLHILES
NLLKAMDSATAPDKIRKLYLYAAHDVTFIPLLMTLGIFDHKWPPFAVDLTMELYQHLESKEW
FVQLYYHGKEQVPRGCPDGLCPLDMFLNAMSVTLSPEKYHALCSQTQVMEVGNEE

Signal sequence:

amino acids 1-23

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 218-222

Casein kinase II phosphorylation site.

amino acids 87-91, 104-108, 320-324

Tyrosine kinase phosphorylation site.

amino acids 280-288

N-myristoylation site.

amino acids 15-21, 117-123, 118-124, 179-185, 240-246, 387-393

Amidation site.

amino acids 216-220

Leucine zipper pattern.

amino acids 10-32

Histidine acid phosphatases phosphohistidine signature.

amino acids 50-65

FIGURE 53

CTCCTCTAACATACTTGCAGCTAAACTAAATATTGCTGCTTGGGACCTCCTCTAGCCT
TAAATTCAGCTCATCACCTCACCTGCCTGGTCATGGCTCTGCTATTCTCCTTGATCCTT
GCCATTGCACCAGACCTGGATTCTAGCGTCTCCATCTGGAGTGCAGCTGGGGCTGGTGGGGGCCT
CCACCGCTGTGAAGGGCGGGTGGAGGTGGAACAGAAAGGCCAGTGGGGCACCGTGTGATG
ACGGCTGGGACATTAAGGACGTGGCTGTGTTGCCGGAGCTGGCTGTGGAGCTGCCAGC
GGAACCCCTAGTGGTATTTGTATGAGCCACCAGCAGAAAAAGAGCAAAAGGTCTCATCCA
ATCAGTCAGTTGCACAGGAACAGAAGATACTGGCTCAGTGTGAGCAAGAAGAAGTTATG
ATTGTTACATGATGAAGATGCTGGGCATCGTGTGAGAACCCAGAGAGCTCTTCTCCCCA
GTCCCAGAGGGTGTCAAGGCTGGCTGACGGCCCTGGCATTGCAAGGGACGCGTGGAAAGTGAA
GCACCAAGAACCAAGTGGTATACCGTGTGCCAGACAGGCTGGAGCCTCCGGCCGAAAGGTGG
TGTGCCGGCAGCTGGATGTGGAGGGCTGTACTGACTCAAAACGCTGCAACAAGCATGCC
TATGGCCGAAAACCCATCTGGCTGAGCCAGATGTCATGCTCAGGACGAGAAGCAACCCTCA
GGATTGCCCTCTGGCCTGGGGAAACAACACCTGCAACCATGATGAAGACACGTGGTCG
AATGTGAAGATCCCTTGACTTGAGACTAGTAGGAGGAGACAACCTCTGCTCTGGCGACTG
GAGGTGCTGCACAAGGGCGTATGGGCTCTGTCATGACAACTGGGAGAAAAGGAGGA
CCAGGTGGTATGCAAGCAACTGGCTGTGGAAAGTCCCTCTCCCTCAGAGACCGGA
AATGCTATGCCCTGGGTTGGCCGATCTGGCTGGATAATGTCGGCTCAGGGAGGAG
CAGTCCCTGGAGCAGTGCAGCACAGATTGGGGTTTCACGACTGCACCCACCAGGAAGA
TGTGGCTGTCATCTGCTCAGTGTAGGTGGCATCATCTAACTGTTGAGTGCCTGAATAGAA
GAAAAACACAGAAGAAGGGAGCATTACTGTCTACATGACTGCATGGATGAACACTGATCT
TCTTCTGCCCTGGACTGGACTTAACTTACTTGGCTGCCCTGATTCTCAGGCCTCAGAGTTGG
ATCAGAACTTACAACATCAGGTCTAGTTCTCAGGCCATCAGACATAGTTGAACTACATCA
CCACCTTCCTATGTCACATTGCACACAGCAGATTCCCAGCCTCCATAATTGTGTAT
CAACTACTAAATACATTCTCACACACACACACACACACACACACACACACACATA
CACCATTTGCTGTTCTGAAGAACTCTGACAAAATACAGATTGGTACTGAAAGAGA
TTCTAGAGGAACGGAATTAAAGGATAAATTCTGAATTGGTATGGGTTCTGAAATTG
GCTCTATAATCTAATTAGATATAAAATTCTGGTAACTTATTTACAATAATAAGATAGCAC
TATGTGTTCAAA

FIGURE 54

MALLFSLILAICTRPGFLASPSGVRLVGGHLRCEGRVEVEQKGQWGTVCDDGWDIKDVAVLC
RELGCGAASGTPSGILYEPPAEKEQKVLIQSVSCTGTEDTLAQCEQEEVYDCSHDEDAGASC
ENPESSFSPVPEGVRLADGPGHCKGRVEVKHQNQWYTVCQTGWSLRAAKVVCRQLGCGRAVL
TQKRCNKHAYGRKPIWLSQMCSGREATLQDCPSGPWGKNTCNHDEDTWVECEDPFDLRLVG
GDNLCSGRLEVLHKGVWGSVCDDNWGEKEDQVVCKQLGCGKSLSPSFRDRKCYGPGVGRIDL
DNVRCSGEEQSLEQCQHRFWGFHDCTHQEDVAVICSV

Signal sequence:

amino acids 1-15

Casein kinase II phosphorylation site.

amino acids 47-51, 97-101, 115-119, 209-213, 214-218, 234-238,
267-271, 294-298, 316-320, 336-340

N-myristoylation site.

amino acids 29-35, 43-49, 66-72, 68-74, 72-78, 98-104, 137-143,
180-186, 263-269, 286-292

Amidation site.

amino acids 196-200

Speract receptor repeated domain signature.

amino acids 29-67, 249-287

FIGURE 55

ACTGCACTCGGTTCTATCGATTGAATTCCCCGGGGATCCTCTAGAGATCCCTCGACCTCGAC
CCACCGCGTCCCGCGACCGTGGCGGACCGTGGGCCGGCTACCAGGAAGAGTCTGCCGAAG
GTGAAGGCCATGGACTTCATCACCTCACAGCCATCCTGCCCTGCTGTTGGCTGCCCTGGG
CGTCTTCGGCCTCTCCGGCTGCTGCAGTGGTGCGCGGAAGGCCTACCTGCGGAATGCTG
TGGTGGTGATCACAGGCGCCACCTCAGGGCTGGCAAAGAATGTGAAAAGTCTTATGCT
GCGGGTGCTAAACTGGTGCTCTGTGGCGGAATGGTGGGCCCTAGAAGAGCTCATCAGAGA
ACTTACCGCTTCTCATGCCACCAAGGTGCAGACACACAAGCCTTACTGGTGACCTTCGACC
TCACAGACTCTGGGCCATAGTTGCAGCAGCAGCTGAGATCCTGCAGTGCTTGGCTATGTC
GACATACTGTCAACAATGCTGGATCAGCTACCGTGGTACCATCATGGACACCACAGTGG
TGTGGACAAGAGGGTCATGGAGACAAACTACTTGGCCAGTTGCTCTAACGAAAGCACTCC
TGCCCTCCATGATCAAGAGGGAGGCAAGGCCACATTGTCGCCATCAGCAGCATCCAGGGCAAG
ATGAGCATTCTTTCGATCAGCATATGCAGCCTCCAAGCACGCAACCCAGGTTTCTT'TGA
CTGTCTGCGTGGAGATGGAACAGTATGAAATTGAGGTGACCGTCATCAGCAGGCCCTACA
TCCACACCAACCTCTGTAAATGCCATCACCGCGGATGGATCTAGGTATGGAGTTATGGAC
ACCACACAGCCCAGGGCGAAGCCCTGTGGAGGTGGCCAGGATGTTCTGCTGCTGTGG
GAAGAAGAAGAAAGATGTGATCCTGGCTGACTTACTGCCTCCTGGCTGTTATCTCGAA
CTCTGGCTCCTGGCTCTTCAGCCTCATGGCCTCCAGGGCCAGAAAAGAGCGGAAATCC
AAGAACTCCTAGACTCTGACCAGCCAGGGCCAGGGCAGAGAAGCAGCACTCTAGGTTGC
TTACTCTACAAGGGACAGTTGCATTGAGACTTTAATGGAGATTGTCTCACAAAGTGG
AAAGACTGAAGAAACACATCTGTGCAGATCTGCTGGCAGAGGACAATCAAAACGACAACA
AGCTTCTTCCCAGGGTGAGGGAAACACTTAAGGAATAAATATGGAGCTGGGTTAACACT
AAAAACTAGAAATAAACATCTCAAACAGTAAAAAAAAAAGGGCGGCCGACTCTAG
AGTCGACCTGCAGAAGCTTGGCCGCCATGGCCCAACTGTTATTGAGCTTATAATGGTTAC

FIGURE 56

MDFITSTAILPLLFGCLGVFGLFRLLQWVRGKAYLRNAVVIITGATSGLGKEAKVFYAAAGA
KLVLCGRNGGALEELIRELTASHATKVQTHKPYLVTFDLTDGAI
VAAAEILQCFGYVDIL
VNNAGISYRGTIMDTTDVDKRVMETNYFGPVALTKALLPSMIKRRQGHIVAISSIQGKMSI
PFRSAYAASKHATQAFFDCLRAEMEQYEIEVTVISPGYIHTNLSVNAITADGSRYGVMDTTT
AQGRSPVEAQDVLA AVGKKKDVLADLLPSLAVYLRTLAPGLFFSLMASRARKERKS KNS

Signal sequence:

amino acids 1-21

Transmembrane domain:

amino acids 104-120, 278-292

N-glycosylation site.

amino acids 228-232

Glycosaminoglycan attachment site.

amino acids 47-51

Casein kinase II phosphorylation site.

amino acids 135-139, 139-143, 253-257

Tyrosine kinase phosphorylation site.

amino acids 145-153, 146-153

N-myristoylation site.

amino acids 44-50, 105-111, 238-244, 242-248, 291-297

Amidation site.

amino acids 265-269

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 6-17

FIGURE 57

CCACCGCGTCCGCTGGTAGATCGAGCAACCCTCTAAAAGCAGTTAGAGTGGTAAAAAA
AAAAAAAAACACACCAAACGCTCGCAGCCACAAAAGGGGATGAAATTCTCTGGACATCCTC
CTGCTTCTCCGTTACTGATCGTCTGCTCCCTAGAGTCCTCGTGAAGCTTTTATTCTAA
GAGGAGAAAATCAGTCACCGGCAGAACATCGTGTGATTACAGGAGCTGGCATGGAATTGGGA
GACTGACTGCCTATGAATTGCTAAACTAAAGCAAGCTGGTCTCTGGATATAATAAG
CATGGACTGGAGGAAACAGCTGCCAATGCAAGGGACTGGGTGCCAAGGTTACACCTTGT
GGTAGACTGCAGCAACCGAGAAGATATTACAGCTCTGCAAAGAAGGTGAAGGCAGAAATTG
GAGATGTTAGTATTAGTAAATAATGCTGGTAGTCTATACATCAGATTGTTGCTACA
CAAGATCCTCAGATTGAAAAGACTTTGAAGTTAATGTAATTGCACATTCTGGACTACAAA
GGCATTCTCCTGCAATGACGAAGAATAACCATGCCATTGTCACTGTGGCTCGCAG
CTGGACATGTCTCGTCCCTTACTGGCTACTGTTCAAGCAAGTTGCTGTTGGA
TTTCATAAAACTTGACAGATGAACACTGGCTGCCTACAAATAACTGGAGTCAAAACAAATG
TCTGTGTCCTAATTCGTAACACTGGCTCATCAAAATCCAAGTACAAGTTGGACCCA
CTCTGGAACCTGAGGAAGTGGTAAACAGGCTGATGCATGGATTCTGACTGAGCAGAAGATG
ATTTTATTCCATCTTCTATAGTTTAAACAACATTGAAAGGATCCTCCTGAGCGTT
CCTGGCAGTTAAAACGAAAATCAGTGTAAAGTTGATGCAGTTATTGGATATAAAATGA
AAGCGCAATAAGCACCTAGTTCTGAAAACGATTACAGGTTAGGTTGATGTCATCTA
ATAGTGCCAGAATTAAATGTTGAACCTCTGTTCTAATTATCCCCATTCTTCAATA
TCATTTTGAGGCTTGGCAGTCTCATTACTACCACTGTTCTTAGCCAAAGCTGATT
ACATATGATATAAACAGAGAAATACCTTAGAGGTGACTTTAAGGAAAATGAAGAAAAGAA
CCAAAATGACTTATTAAAATAATTCCAAGATTGTTGCTCACCTGAAGGCTTGCAA
AATTTGTACCATACCGTTATTAAACATATATTGTTGATTGACTTAAATTGTTG
ATAATTGTGTTCTTTCTGTTCTACATAAAATCAGAAACTCAAGCTCTCTAAATAAAA
TGAAGGACTATCTAGTGGTATTTCACAATGAATATCATGAACCTCAATGGTAGGTTTC
ATCCTACCCATTGCCACTCTGTTCTGAGAGATAACCTCACATTCAATGCCAAACATTCT
GCACAGGGAAGCTAGAGGTGGATACAGTGTGCAAGTATAAAAGCATCACTGGATTAAAG
GAGAATTGAGAGAATGTACCCACAAATGGCAGCAATAATAATGGATCACACTTAAAAAA
AA
AAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 58

MKFLLDILLPLLLIVCSLESFVKLFIPKRRKSVTGEIVLITGAGHIGRLTAYEFAKLKSK
LVLWDINKHGLEETAAKCKGLGAKVHTFVVDCSNREDIYSSAKVKAEIGDVSILVNNAGVV
YTSDFLFAQDPQIEKTFEVNVLAHFWTTKAFLPAMTKNNHGHIVTVASAAGHVSVPFLLAYC
SSKFAAVGFHKTTLTDELAALQITGVKTTCLCPNFVNTGFIKNPSTSLGPTLEPEEVVNRLMH
GILTEQKMIFIPSSIAFLTTLERILPERFLAVLKRKISVKFDAVIGYKMKAQ

Signal sequence:

amino acids 1-19

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 30-34, 283-287

Casein kinase II phosphorylation site.

amino acids 52-56, 95-99, 198-202, 267-271

N-myristoylation site.

amino acids 43-49, 72-78, 122-128, 210-216

FIGURE 59

CCACCGCGTCCGCGGACCGCGTGGTCGACTAGTTCTAGATCGCGAGCGGCCGCCGGCTC
AGGGAGGAGCACCGACTCGCGCCGACCCCTGAGAGATGGTGGTGCCATGTGGAAGGTGATTG
TTTCGCTGGTCTGTTGATGCCTGGCCCTGTGATGGGCTGTTCGCTCCCTATACAGAAGT
GTTTCCATGCCACCTAAGGGAGACTCAGGACAGCCATTATTCTCACCCCTTACATTGAAGC
TGGGAAGATCCAAAAGGAAGAGAATTGAGTTGGTCGGCCCTTCCAGGACTGAACATGA
AGAGTTATGCCGGCTTCCTCACCGTGATAAGACTTACAACAGCAACCTCTTCTGGTTC
TTCCCAGCTCAGATACAGCCAGAAGATGCCAGTAGTTCTCTGGCTACAGGGTGGGCCGG
AGGTTCATCCATGTTGGACTCTTGGAACATGGCCCTATGTTGTCACAAGTAACATGA
CCTTGCCTGACAGAGACTTCCCTGGACCACAACGCTCTCCATGCTTACATTGACAATCCA
GTGGGCACAGGCTTCAGTTTACTGATGATACCCACGGATATGCAATGAGGACGATGT
AGCACGGGATTATACAGTGCACTAATTCAAGTTTCCAGATATTCTGAATATAAAAATA
ATGACTTTATGTCACTGGGAGTCTTATGCAGGGAAATATGTGCCAGCCATTGCACACCTC
ATCCATTCCCTCAACCCTGTGAGAGAGGTGAAGATCAACCTGAACGGAATTGCTATTGGAGA
TGGATATTCTGATCCGAATCAATTATAGGGGCTATGCAGAATTCTGTACCAAATTGGCT
TGTTGGATGAGAAGCAAAAAAGTACTTCCAGAAGCAGTGCCTGAATGCATAGAACACATC
AGGAAGCAGAACTGGTTGAGGCCTTGAATACTGGATAAACTACTAGATGGCGACTAAC
AAGTGCCTTCTTACTTCCAGAATGTTACAGGATGTAGTAATTACTATAACTTTGCGGT
GCACGGAACCTGAGGATCAGCTTACTATGTGAAATTGGTCACTCCCAGAGGTGAGACAA
GCCATCCACGTGGGAATCAGACTTTAATGATGGAACTATAGTTGAAAAGTACTTGCAGA
AGATAACAGTACAGTCAGTTAACGCATGGTTAAGTGAATCATGAATAATTATAAGGTTCTGA
TCTACAATGGCCAACGGACATCATCGTGGCAGCTGCCCTGACAGAGCGCTCTTGATGGC
ATGGACTGGAAAGGATCCAGGAATACAAGAAGGCAGAAAAAGTTGGAAGATCTTAA
ATCTGACAGTGAAGTGGCTGGTTACATCCGGCAAGCGGGTGAATTCCATCAGGTAATTATTC
GAGGTGGAGGACATATTTCACCTATGACCAGCCTGTGAGAGCTTGACATGATTAATCGA
TTCATTATGGAAAAGGATGGATCTTATGTTGGATAAAACTACCTTCCAAAAGAGAACAT
CAGAGGTTTCATTGCTAAAAGAAAATCGAAAAACAGAAAATGTCATAGGAATAAAAAAA
TTATCTTTCATATCTGCAAGATTCTTCAATAAAAATTATCCTGAAACAAAGTGCAG
TTTGTGTTGGGGGAGATGTTACTACAAAATTACATGAGTACATGAGTAAGAATTACA
TTATTTAAAGGATGAAAGGTATGGATGTGACACTGAGACAAGATGTATAAATGA
AATTTAGGGTCTTGAATAGGAAGTTAATTCTTAAGAGTAAGTGAAGAAGTGCAGTTG
TAACAAACAAAGCTGAACATCTTCTGCCAATAACAGAAGTTGGCATGCCGTGAAGGT
GTTTGGAAATATTATGGATAAGAATAGCTCAATTATCCAAATAATGGATGAAGCTATAA
TAGTTTGGGAAAAGATTCTCAAATGTATAAAGTCTTAGAACAAAGAACATTCTTGAAATA
AAAATATTATATAAAAGTAAAAAA

FIGURE 60

MVGAMWKVIVSLVLLMPGPCDGLFRSLYRSVSMPKGDSQPLFLTPYIEAGKIQKGRELSL
VGPFPGLNMKSYAGFLTVNKTYNSNLFFWFFPAQIQPEDAPVVLWLQGGPGGSSMFGLFVEH
GPYVVTSNMTLRDRDFPWTTLSMLYIDNPVGTGFSFTDDTHGYAVNEDDVARDLYSALIQF
FQIFPEYKNNDFYVTGESYAGKYVPAIAHLIHSLNPVREVKINLNGIAIGDGYSDPESIIGG
YAEFLYQIGLLDEKQKKYFQKQCHECIEHIRKQNWFEAFEILDKLLDGLTSDPSYFQNVTG
CSNYYNFLRCTEPEDQLYYVKFLSLPEVRQAIHVGNQTFNDGTIVEKYLREDTVQSVKPWLT
EIMNNYKVLIYNGQLDIIVAALTERSLMGMDWKGSQEYKKAEKVVKIFKSDSEVAGYIRO
AGDFHQVIIRGGGHILPYDQPLRAFDMINRFIYGKGWDPYVG

Signal sequence:

amino acids 1-22

N-glycosylation site.

amino acids 81-85, 132-136, 307-311, 346-350

Casein kinase II phosphorylation site.

amino acids 134-138, 160-164, 240-244, 321-325, 334-338, 348-352,
353-357, 424-428

Tyrosine kinase phosphorylation site.

amino acids 423-432

N-myristoylation site.

amino acids 22-28, 110-116, 156-162, 232-238

Serine carboxypeptidases, serine active site.

amino acids 200-208

Crystallins beta and gamma 'Greek key' motif signature.

amino acids 375-391

FIGURE 61

CGAGGGCTTTCCGGCTCCGAATGGCACATGTGGATCCCAGTCTTGGCTACAACAT
 TTTCCCTTCCTAACAAAGTCTAACAGCTGTTCAACAGCTAGTATCAGGGTTCTTCTT
 GCTGGAGAAGAAAGGGCTGAGGGCAGAGCAGGGCACTCTCACTCAGGGTGACCAGCTCCTT
 CCTCTGTGGATAACAGAGCATGAGAAAGTGAAGAGATGCAGCGAGTGAGGTGATGAAAG
 TCTAAATAGGAAGGAATTGTGCAATATCAGACTCTGGGAGCAGTTGACCTGGAGAGC
 CTGGGGAGGGCTGCTAACAGCTTCAAAAAACAGGAGCGACTCCACTGGCTGGGAT
 AAGACGTGCCGGTAGGATAGGGAAGACTGGGTTAGTCTTAATATCAAATTGACTGGCTGG
 TGAACTTCAACAGCCTTAAACCTCTGGGAGATGAAAACGATGGCTTAAGGGGCCAGAAA
 TAGAGATGCTTGAAATAAAATTAAAAAGCAAGTATTTATAGCATAAAGGCTAGA
 GACCAAAATAGATAACAGGATTCCCTGAACATCCTAAGAGGGAGAAAGTATGTTAAAAATA
 GAAAAACCAAAATGCAGAAGGGAGGAGACTCACAGAGCTAACACCAGGATGGGACCTGGTC
 AGGCCAGCCTTTGCTCCTCCGAAATTATTTGGTCTGACCCTCTGCCTTGTGTTT
 GCAGAATCATGTGAGGGCAACCGGGGAAGGTGGAGCAGATGAGCACACACAGGAGCCGTCT
 CCTCACCGCCGCCCTCTCAGCATGGAACAGAGGCAGCCCTGGCCCGGGCCCTGGAGGTGG
 ACAGCCGCTGTGGCTCTCAGTGGCTGGGTCTGGGCTGGCTGGCCCCCAGCAGCCGGC
 ATGCCTCAGTCAGCACCTCCACTCTGAGAATCGTACTGGACCTCAACCACCTGACCGT
 CCACCAAGGGACGGGGCCGTCTATGTGGGGCATCAACGGGTCTATAAGCTGACAGGCA
 ACCTGACCATCCAGGTGGCTATAAGACAGGGCCAGAAGAGGACAACAAGTCTCGTACCCG
 CCCCTCATCGTGAGCCCTGCAGCGAAGTGTCTACCCCTACCAACAATGTCAACAAGCTGCT
 CATCATTGACTACTCTGAGAACCGCCGTCTGGCTGTGGAGCCTCTACCAGGGGTCTGCA
 AGCTGCTGCGGCTGGATGACCTCTTCATCCTGGTGGAGCCATCCCACAAGAAGGAGCACTAC
 CTGTCAGTGTCAACAAGACAGGGCACCAGTACGGGTGATTGTGCGCTGTGAGGTGAGGA
 TGGCAAGCTCTCATGGCACGGCTGTGGATGGGAAGCAGGATTACTTCCCACCCCTGTCCA
 GCCGGAAGCTGCCCGAGACCCCTGAGTCCTCAGCCATGTCGACTATGAGCTACACAGCGAT
 TTTGTCCTCTCATCAAGATCCCTCAGACACCCCTGGCCCTGGTCTCCACTTTGACAT
 CTTCTACATCTACGGCTTGCTAGTGGGGCTTGTCTACTTCTCACTGTCAGCCCGAGA
 CCCCTGAGGGTGTGGCATCAACTCCGCTGGAGACCTCTACACCTCACGCATCGTGC
 CTCTGCAAGGATGACCCCAAGTCCACTCATCGTGTCCCTGCCCTCGGCTGCACCCGGC
 CGGGTGGAAATACCGCCTCCTGCAGGCTGCTACCTGGCCAAGCCTGGGACTCACTGGCC
 AGGCCTCAATATCACCAGCCAGGACGATGTACTCTTGCCATCTCTCAAAGGGCAGAAG
 CAGTATCACCACCGCCGATGACTCTGCCCTGTGCGCTTCCATCGGACATCAACTT
 GCAGATCAAGGAGCGCCTGCAGTCTGCTACCAGGGCGAGGGCAACCTGGAGCTCAACTGGC
 TGCTGGGAAGGACGTCCAGTGCACGAAGGCCCTGTCCCCATCGATGATAACTCTGTGGA
 CTGGACATCAACCAGCCCTGGGAGGCTCAACTCCAGTGGAGGGCTGACCCGTACACCAC
 CAGCAGGGACCGCATGACCTCTGTGCCCTCACGTTACAACGGCTACAGCGTGGTTTTG
 TGGGACTAAGAGTGGCAAGCTAAAAAGGTAAGAGTCTATGAGTTAGATGCTCAATGCC
 ATTCACTCCTCAGCAAAGAGTCCCTTGGAGGTAGCTATTGGTGGAGATTAACTATAG
 GCAACTTATTTCTTGGGAACAAAGGTGAAATGGGAGGTAAGAAGGGTTAATTGTG
 ACTTAGCTTAGCTACTTCCAGCCATCAGTCATTGGTATGTAAGGAATGCAAGCGTA
 TTCAATATTCCAAACTTAAGAAAAACTTAAGAAGGTACATCTGCAAAAGCAAA

FIGURE 62

MGTLGQASLFAPPGNYFWSDHSALCFAESCEGQPGKVEQMSTHRSRLLTAAPLSMEQRQPWP
RALEVDSRSVVLSSVVWVLLAPPAAGMPQFSTFHSENRDWTFNHLTVHQGTGAVYVGAINRV
YKLTGNLTIQVAHKTGPEEDNKSRYPPLIVQPCSEVLTLTNNVNKLLIIDYSENRLLAGSL
YQGVCKLLRLDDLFILVEPSHKKEHYLSSVNKTGTMGVIVRSEGEDGKLFIGTAVDGKQDY
FPTLSSRKLPRDPESSAMLDYELHSDFVSSLIKIPSDTLALVSHFDIFYIYGFASGGFVYFL
TVQPETPEGVAINSAGDLFYTSRIVRLCKDDPKFHSYVSLPFGCTRAGVEYRLLQAAYLAKP
GDSLAQAFNITSQDDVLFAIFSKGQKQYHHPPDDSALCAFPIRAINLQIKERLQSCYQGEGN
LELNWLLGKDVQCTKAPVPIDDNFCGLDINQPLGGSTPVEGLTLYTTSRDRMTSVASYVYNG
YSVVFVGTSGKLKKVRVYEFRCSNAIHLLSKESLLEGSYWWRFNYRQLYFLGEQR

Signal sequence:

amino acids 1-32

Transmembrane domain:

amino acids 71-87

N-glycosylation site.

amino acids 130-134, 145-149, 217-221, 381-385

Casein kinase II phosphorylation site.

amino acids 139-143, 229-233, 240-244, 291-295, 324-328, 383-387,
384-388, 471-475, 481-485, 530-534

N-myristoylation site.

amino acids 220-226, 319-325, 353-359, 460-466, 503-509

FIGURE 63

AGGCTCCCGCGCGCGCTGAGTGGACTGGAGTGGGACCCGGGCCCCCGCTAGAGAACACGCGATGACCA
 CGTGGAGCCTCCGGCGAGGCCGGCCCGCACGCTGGACTCTGCTGCTGGCTCTGGGCTCCTGGCTCC
 GCAGGCTGGACTGGAGCACCCCTGGCCCTCGCGCTCCGCATCGACAGCTGGGCTGCAGGCCAAGGGCTGGA
 ACTTCATGCTGGAGGATTCCACCTCTGGATCTCGGGGCTCCATCCACTATTTCCGTGCCCAGGGAGTACT
 GGAGGGACCGCTGCTGAAGATGAAGGCTGTGGCTTGAACACCCCTACACCACCTATGTTCCGTGAAACCTGCATG
 AGCCAGAAAGAGGCAAATTGACTTCTCTGGAACCTGGACCTGGAGGCCCTCGTCTGATGGCCAGAGATCG
 GGCTGTGGGTGATTCTGCTCCAGGCCCTACATCTGCAGTGGAGATGGACCTCGGGGCTTGCCCAGCTGGCTAC
 TCCAAGACCCCTGGCATGAGGCTGAGGACAACTTACAAGGCTCACCGAAGCAGTGGACCTTATTTGACCACC
 TGATGTCCAGGGTGGTGCACCTCCAGTACAAGCTGGGGACCTATCATTGCCGTGAGGTGAGAATGAATATG
 GTTCCCTATAAAAGACCCGCATACATGCCCTACGTCAAGAAGGCACTGGAGGACCGTGGCATTGTGAACTGC
 TCCCTGACTTCAGACAACAAGGATGGGCTGAGCAAGGGATTGTCCAGGGAGTCTTGGCCACCATCAACTTGCACT
 CAACACAGGACTGCAGCTACTGACCACCTTCTCTCAACGTCCAGGGACTCAGGCCAAGATGGTGTGGAGT
 ACTGGACGGGGTGGTTGACTCGTGGGGAGGCCCTCACAAATATCTGGATTCTCTGAGGTTTGAAAACCGTGT
 CTGCCATTGTGGACGCCGCTCCATCAACCTCTACATGTTCCACGGAGGCCAACACTTGGCTTATGAAATG
 GAGCCATGCACCTCCATGACTACAAGTCAGATGTCAACAGCTATGACTATGATGCTGTGCTGACAGAACGGCG
 ATTACACGGCCAAGTACATGAAGCTCGAGACTTCTCGCTCCATCTCAGGCATCCCTCCCTCCCCCACCTG
 ACCTTCTTCCAAGATGCCGTATGCCCTAACGCCAGTCTTGTACCTGTCTGTGGGACGCCCTCAAGTACC
 TGGGGGAGCCAATCAAGTCTGAAAAGCCCATCAACATGGAGAACCTGCCAGTCATGGGGAAATGGACAGTCCT
 TCGGGTACATTCTCATGAGACCAGCATCACCTCGTCTGCCATCCTCAGTGGCCACGTGCATGATGGGGCAGG
 TGTGTTGAAACAGTATCCATAGGATTCTGGACTACAAGACAACGAAGATTGCTGCCCCCTGATCCAGGGTT
 ACACCGCTGCTGAGGATCTTGGTGGAGAATGTGGCGAGTCACATGGGGAGAATTATGATGACCAAGCGCAAAG
 GCTTAATTGAAATCTCTATCTGAATGATTCAACCTTGTGAAACATGGGGAAATGGACAGTCCT
 GCTTCTTCAGAGGTTCGGCCTGGACAAATGGNTCCCTCCAGAAACACCCACATTACCTGCTTTCTTGTG
 GTAGCTTGCCATCAGCTCCACGCCCTGTGACACCTTCTGAAGCTGGAGGGCTGGAGAAGGGGTTGATTCA
 TCAATGCCAGAACCTTGGACGTACTGGACCATGGACCCAGAAGACGCTTACCTCCAGGTCCCTGGTGA
 GCAGCGGAATCAACCAAGTCATGTTTGAGGAGACGATGGGGCCCTGCATTACAGTTCACGGAAACCCCC
 ACCTGGGAGGAACCAGTACATTAAGTGAGGGTGGCACCCCTCCTGCTGGTGCAGTGGAGACTGCC
 CTCTTGACCTGAAGCTGGTGGCTGCTGCCACCCCTACTGCAAAGCATTAAAGTGCACAGG
 ACTGGGGCTACAGTCTGCCCTGTCTCAGCTCAAAACCTAACGCTGCAGGGAAAGGTGGGATGGCTGGG
 TGGTTTGTGATGATGGCTTCTACAGCCCTGCTCTTGTGCCAGGGCTGTGGCTGTCTAGGGTGGAGC
 AGCTAATCAGATGCCAGGCTTGGCCCTCAGAAAAGTGCAGAACACGTCACAGGCC
 TGCGAGCATCTGCTGACTCAGGCGTGTCTTGCTGGTCTGGGAGGCTTGGGACATCCCTCATGGCCCC
 TTTATCCCCGAAATCTGGGTGTCACCAGTGTAGAGGGTGGGAAGGGGTTCTCACCTGAGCTGACTTGT
 CTTCCTTCACAACCTCTGAGCCTTCTTGGGATTCTGAAAGGAACCTGGCGTGAAGAAACATGTGACTTCC
 TCCCTTCCACTCGCTGCTCCACAGGGTGACAGGCTGGGCTGGAGAAACAGAAATCCTCACCC
 CAAGTTAGCAGGTGCTCTGGTGTGAGTGGAGGACATGTGAGTCTGGCAGAAGCCATGGCC
 CATCCAGGGAGGAGGACAGAACGGCCAGCTCACATGTGAGTCTGGCAGAACGCCATGGCC
 AGGGAGGAGGACAGAACGGCCAGCTCACATGTGAGTCTGGCAGAACGCCATGGCC
 GGAGGACAGAACGGCCAGCTCACATGTGAGTCTGGCAGAACGCCATGGCC
 ACAGAACGGCCAGCTCACATGTGAGTCTGGCAGAACGCCATGGCC
 GAAGTGTGTCAGTCCGCAAGTCCGCAATTGAGCCTTGTCTGGG
 CCCAGCCAAACACCTGGCTTGGGACTGTGCTG
 GTTGAGTAAAGCTATAACCTTGAAATCACA

FIGURE 64

MTTWSLRRR PARTLGLLLL VVVLGFLV LRR LDWSTL VPLRLR HQL GLQAK GWN F MLED STFW
IFGGSIHYFRVPREYWRDRLLKMKACGLNTLTTYVPWNLH EPERGKFDFSGNL DLEAF VLMA
AEIGLWVILRPGPYICSEMDLGGLPSWLLQDPGMRLRTTYKGFT EAVDLYFDHLM SRV VPLQ
YKRGGPIIAVQVENEYGSYNKDPAYMPYVKKALEDRGIV ELLLTS DNKD GLSKGIV QGV L AT
INLQSTHELQ LTTFLFNVQGTQPKMVMEYWTGWFDSWGGPHNILD SSEVLKTVSAIVDAGS
SINLYMFHGGTNFGFMNGAMHFHDYKSDVTSYDYDAVLTEAGDYTA KYMKL RDFFGSISGIP
LPPPPD L LPKMPYEP LTPVLYLSLWDALKYLGEPIKSEKPINMENLPVNGGNQSF GYI LY
TSITSSGILSGHVHD RGQVFVN T VSIGFLDYKTTKIAVPLI QGYTVLRILVENRGRVN YGEN
IDDQRKGLIGNLYLNDSP LKNFRIYSLDMKKSFFQRFGLDKWXSLP EPTLPAFFLGSLSIS
STPCDTFLKLEGWEKGVVFINGQNLGRYWNIGPQKTLYLPGPWLS SGINQVIVFEETMAGPA
LQFTETPHLGRNQYIK

Signal sequence:

amino acids 1-27

Casein kinase II phosphorylation site.

amino acids 141-118, 253-257, 340-344, 395-399, 540-544, 560-564

N-myristoylation site.amino acids 146-152, 236-242, 240-246, 244-250, 287-293, 309-315,
320-326, 366-372, 423-429, 425-431, 441-447, 503-509, 580-586

FIGURE 65

GGGGACGCGGAGCTGAGAGGCTCCGGCTAGCTAGGTGTAGGGTGGACGGTCCCAGGACC
 CTGGTAGGGTTCTCTACTTGGCCTCGTGGGGTCAAGACGCAGCACCTACGCCAAGG
 GGAGCAAAGCCGGCTCGGCCCGAGGCCCGAGCACCTCCATCTCCAATGTTGGAGGAATC
 CGACACGTACGGTCTGTCCCGTCTCAGACTAGAGGAGCGCTGTAAACGCCATGGCTCCC
 AAGAAGCTGCTGCCTCGTCCCTGCTGCTGCCGCTCAGCTGACGCTACTGCTGCCCA
 GGCAGACACTCGTCGTTCTGAGGGATAGGGTCACTGACCGGTTCTCCTAGACGGGGCC
 CGTCCGCTATGTGTCTGGCAGCCTGCACTACTTCGGGTACCGCGGGTCTTGGGCCAC
 CGGTTTGAGATGCGATGGAGCGGCCCAACGCCATACAGTTTATGTGCCCTGGAACTA
 CCACGAGGCCACAGCCTGGGGTCTATAACTTAAATGGCAGCCGGACCTCATTGCCCTTCTGA
 ATGAGGCAGCTAGCGAACCTGTTGGTCATACTGAGACCAAGGACCTACATCTGTGCAGAG
 TGGGAGATGGGGGTCTCCATCCTGGTCTCGAAAACCTGAAATTCAAGAACCTC
 AGATCCAGACTTCCCTGCCAGTGGACTCCTGGTCAAGGTCTGCTGCCAAGATATATC
 CATGGCTTATCACAAATGGGGCAACATCATTAGCATTAGGTGGAGAATGAATATGGTAGC
 TACAGAGCCTGTGACTTCAGCTACATGAGGCACCTGGCTGGCTCTCCGTGCACTGCTAGG
 AGAAAAGATCTTGTCTTACACAGATGGCCTGAAGGACTCAAGTGTGGCTCCCTCCGG
 GACTCTATACCACTGTAGATTGGCCAGCTGACAACATGACCAAAATCTTACCTGCTT
 CGGAAGTATGAACCCATGGCATTGGTAAACTCTGAGTACTACACAGGCTGGCTGGATTA
 CTGGGCCAGAATCACTCCACACGGTCTGTCAAGCTGTAACCAAAAGGACTAGAGAACATGC
 TCAAGTTGGAGCAGTGTGAACATGTACATGTTCCATGGAGGTACCAACTTGGATATTGG
 AATGGTGCCTGATAAGAAGGGACGCTCCTCCGATTACTACCAGCTATGACTATGATGCACC
 TATATCTGAAGCAGGGGACCCCACACCTAACGTTTGCTCTCGAGATGTCATCAGCAAGT
 TCCAGGAAGTTCTTGGACCTTACCTCCCCGAGCCCCAACAGATGATGCTTGGACCTGTG
 ACTCTGCACCTGGTGGCATTACTGGCTTCTAGACTGCTTGCCTGGCCAT
 TCATTCAATCTTGCCTGACCTTGGAGCTGCAAGCAGGACCATGGCTCATGTTGACC
 GAACCTATATGACCCATACCACTTTGAGCCAACACCATTCTGGGTGCCAATAATGGAGTC
 CATGACCGTGCCTATGTGATGGTGGATGGGGTGTCCAGGGTGTGGAGCGAAATATGAG
 AGACAAACTATTTGACGGGAAACTGGGTCAAACCTGGATATCTTGGTGGAGAACATGG
 GGAGGCTCAGTTGGTCTAACAGCAGTGAACAGGCCTGTTGAAGCCACCAATTCTG
 GGGCAACAAATCCTTACCCAGTGGATGATGTTCCCTCTGAAAATTGATAACCTGTGAAGTG
 GTGGTTCCCTCCAGTGCCTAACGGCCATATCCTCAAGCTCCTCTGGCCCCACATTCT
 ACTCCAAAACATTCCAATTAGGCTCAGTTGGGACACATTCTATATCTACCTGGATGG
 ACCAAGGGCCAAGTCTGGATCAATGGTTAACCTGGCCGGTACTGGACAAAGCAGGGCC
 ACAACAGACCCCTACGTGCCAAGATTCTGCTGTTCCCTAGGGGAGCCCTAACAAAATTA
 CATTGCTGGAACCTAGAAGATGTACCTCTCCAGCCCCAACGCTAACATTGGATAAGCCTATC
 CTCATAGCACTAGTACTTGCACAGGACACATATCAATTCCCTTCAGCTGATAACTGAG
 TGCCTCTGAACCAATGGAGTTAAGTGGCACTGAAGGTAGGCCGGCATGGTGGCTCATGC
 CTGTAATCCAGCACTTGGGAGGCTGAGACGGGTGGATTACCTGAGGTCAAGGACTTCAAGA
 CCAGCCTGGCCAACATGGTAAACCCGTCTCACTAAAAATACAAAATTAGCCGGCGTG
 ATGGTGGGCACCTTAATCCAGCTACTTGGGAGGCTGAGGGCAGGGAGAATTGCTTGAATCC
 AGGAGGCAGAGGTTGCAGTGAGTGGAGGTTGTACCACTGCACTCCAGCCTGGCTGACAGTGA
 GACACTCCATCTCAAAAAAA

FIGURE 66

MAPKKLSCLRSLLLPLSLTLLPQADTRSFVVDRGHDRFLLDGAPFRYVSGSLHYFRVPRVL
WADRLLKMRWSGLNAIQFYVPWNYHEPQPGVYNFNGSRDLIAFLNEAALANLLVILRPGPYI
CAEWEMGGLPSWLLRKPEIHLRTSDPDFLAAVDSWFKVLLPKIYPWLYHNGGNIISIQVENE
YGSYRACDFSYMRHLAGLFRALLGEKILLFTTDGPEGLKCGSLRGLYTTVDFGPADNMTKIF
TLLRKYEPHGPLVNSEYYTGWLDYWGQNHSTRVSATKGLENMLKLGASVNMYMFHGGTNF
GYWNGADKKGRFLPITTSYDYDAPISEAGDPTPKLFALRDVISKFQEVPPLGPLPPPSPKMMI
GPVTLHLVGHLLAFLDLLCPRGPIHSILPMTFEAVKQDHGFLYRTYMTHTIFEPPTFWVPN
NGVHDRAYVMVDGVFQGVVERNMRDKLFLTGKLGSKLDILVENMGRLSFGSNSSDFKGLLKP
PILGQTILTQWMMFPLKIDNLVKWWFPLQLPKWPYPQAPSGPTFYSKTFPILGSGVDTFLYL
PGWTKGQWINGFNLGRYWTQGPQQTLYVPRFLFPRGALNKITLLELEDVPLQPQVQFLD
KPILNSTSTLHRTHINSLSADTLSASEPMELSGH

Signal sequence:

amino acids 1-27

N-glycosylation site.

amino acids 97-101, 243-247, 276-280, 486-490, 625-629

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 4-8

Casein kinase II phosphorylation site.

amino acids 148-152, 234-238, 327-331, 423-427, 469-473, 550-554,
603-607, 644-648

Tyrosine kinase phosphorylation site.

amino acids 191-198

N-myristoylation site.

amino acids 131-137, 176-182, 188-194, 203-209, 223-229, 227-233,
231-237, 274-280, 296-300, 307-313, 447-453, 484-490

FIGURE 67

GCTTTGAACACGTCTGCAAGCCAAAGTTGAGCATCTGATTGGTTATGAGGTATTCAGTGC
 ACCCACAATATGGCTTACATGTTGAAAAAGCTTCTCATCAGTTACATATCCATTATTTGTGT
 TTATGGCTTATCTGCCTCTACACTCTCTGGTTATTCAAGGATACCTTGAAGGAATATT
 CTTTCGAAAAAGTCAGAGAAGAGAGCAGTTAGTGACATTCCAGATGTCAAAAACGATTT
 GCGTTCCCTCTTCACATGGTAGACCAGTATGACCAGCTATATTCCAAGCGTTGGTGTGTT
 CTTGTCAGAAGTTAGTGAAAATAACTAGGGAAATTAGTTGAACCATGAGTGGACATTG
 AAAAACTCAGGCAGCACATTACGCAACGCCAGGACAAGCAGGAGTTGCATCTGTTCATG
 CTGTCGGGGGTGCCGATGCTGTCTTGACCTCACAGACCTGGATGTGCTAAAGCTTGAAC
 AATTCCAGAAGCTAAAATTCTGCTAACAGATTCTCAAATGACTAACCTCCAAGAGCTCCACC
 TCTGCCACTGCCCTGCAAAAGTTAACAGACTGCTTTAGCTTCTCGCGATCACTTGAGA
 TGCCCTCACGTGAAGTCACTGATGTGGCTGAAATTCTGCCTGGGTGATTGCTCAAAA
 CCTTCGAGAGTTGTACTTAATAGGCAATTGAACTCTGAAAACAATAAGATGATAGGACTTG
 AATCTCTCCGAGAGTTGCGGCACCTTAAGATTCTCCACGTGAAGAGCAATTGACCAAAGTT
 CCCTCCAACATTACAGATGTGGCTCCACATCTAACAAAGTTAGTCATTACATAATGACGGCAC
 TAAACTCTGGTACTGAACAGCCTAACAGAAAATGATGAATGTCGCTGAGCTGGAACCTCCAGA
 ACTGTGAGCTAGAGAGAATCCCACATGCTATTTCAGCCTCTCTAACAGGAACGGAT
 TTAAAGTCCAATAACATTGCGACAATTGAGGAAATCATCAGTTCCAGCATTAAAACGACT
 GACTTGTAAATTATGGATAACAAAATTGTTACTATTCCCTCCCTCTATTACCCATGTCA
 AAAACTTGGAGTCACTTATTCTTAACAAACAAGCTCGAACCTTACCACTGGCAGTATT
 AGTTTACAGAAAATCAGATGCTTAGATGTGAGCTACAACAACATTCAATGATTCCAATAGA
 AATAGGATTGCTTCAGAACCTGCAGCATTGATATCACTGGGAACAAAGTGGACATTCTGC
 CAAAACAATTGTTAAATGCATAAAAGTTGAGGACTTGAATCTGGACAGAACTGCATCACC
 TCACTCCCAGAGAAAAGTTGGTCAGCTCTCCAGCTCACTCAGCTGGAGCTGAAGGGGAAC
 CTTGGACCGCCTGCCAGGCCAGCTGGGCCAGTGTGGATGCTCAAGAAAAGCGGGCTTGGT
 TGGAAGATCACCTTTGATAACCTGCCACTCGAACGTCAAAGAGGGATTGAATCAAGACATA
 AATATTCCCTTGCAAATGGGATTTAAACTAAGATAATATGCACAGTGATGTGCAGGAAC
 AACTTCCTAGATTGCAAGTGCTACGTACAAGTTATTACAAGATAATGCATTAGGAGTAG
 ATACATCTTTAAAATAACAGAGAGGATGCATAGAAGGCTGATAGAAGACATAACTGAAT
 GTTCAATGTTGTAGGGTTAAAGTCATTCAAATCATTTTTTTCTTTGGGG
 AAAGGGAAGGAAAATTATAACTAAATCTTGGTTCTTTAAATTGTTGTAACCTGGAT
 GCTGCCGCTACTGAATGTTACAAATTGCTTGCCCTGCTAAAGTAAATGATTAAATTGACATT
 TTCTTACTAAAAAAAAAAAAAA

FIGURE 68

MAYMLKKLLISYISIICVYGFICLYTLFWLFRIPLKEYSFEKVREESSFSDIPDVKNDFAFL
LHMVDQYDQLYSKRGVFLSEVSENKLREISLNHEWTFEKLHQHISRNAQDKQELHLFMLS
VPDAVFDLTDLDVLKLELIPAKIPIAKISQMTNLQELHLCHCPAKVEQTAFLSFLRDHLRCLH
VKFTDVAEIPAWVYLLKNLRELYLIGNLNSENNKMICLQELRHLKILHVKSNLTKVPSN
ITDVAPHLTKLVIHNDGTKLLVLNSLKKMMNVAELELQNCELERIPHAIFSLSNLQELDLKS
NNIRTIEEIIISFQHLKRLTCLKLWHNKIVTIPPSITHVKNLESLYFSNNKLES LPVAVFSLQ
KLRCLDVSYNNISMIPIEIGLLQNLQHLHITGNKVDILPKQLFKCIKLRTLNLGQNCITS LP
EKVGQLSQLTQLELKGNCLDRLPQLGQCRMLKKSGLVVEDHLFDTLPLEVKEALNQDINIP
FANGI

Signal sequence:

amino acids 1-20

N-glycosylation site.

amino acids 241-245, 248-252, 383-387

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 326-330

Casein kinase II phosphorylation site.

amino acids 48-52, 133-137, 226-230, 315-319, 432-436, 444-448

Tyrosine kinase phosphorylation site.

amino acids 349-355, 375-381

N-myristoylation site.

amino acids 78-84, 124-130, 212-218, 392-398

FIGURE 69

CCCACGCGTCCGGCCTCTCTGGACTTGATTCATTCCATTCTGGAGACACTTGTGTTTACACACATAAGGAT
 TTTTTCCATCTCTGGCCAGCTTGGATCTAGGCCCTGGGAAGACATTGTGTTTACACACATAAGGAT
 CTGTGTTGGGTTCTTCTCCCTGGACATTGGCATTGCTTAGTGGTTGTGGGGAGGGAGACCACGTGG
 GCTCAGTGCCTGCTGCACTTATCTGCCTAGGTACATCGAAGTCTTGACCTCCATACAGTGATTATGCCTGTC
 ATCGCTGGGTATCTGGCGGCCCTGCTCTGCTGATAGTGTGCTCTGTCTTACTTCAAAATACACAAC
 GCGCTAAAAGCTGCAAAGAACCTGAAGCTGGCTGAAAAATACAACCCAGACAAGGTGTGGTGGGCAAG
 AACAGCCAGGCCAAACCAATTGCCACGGAGCTTGTCTGCCCTGCAGTGTGAAGGGATATAGAATGTGTGCC
 AGTTTGATTCCCTGCCACCTTGTGTTGCGACATAAATGAGGGCCTGTGAGTTAGGAAAGGCTCCCTCTCAA
 GCAGAGCCCTGAAGACTTCAATGATGTCATGAGGCCACCTGTTGTGATGTGAGGGCACAGAAGAAGGCACAG
 CTCCCCATCAGTTATGGAAAATACTCAGTGCCTGCTGGGAACAGCTGCTGGAGATCCCTACAGAGAGCTTC
 CACTGGGGCAACCCCTCCAGGAAGGAGTTGGGAGAGAGAACCCACTGTGGGGAAATGCTGATAAAACCAAGTCA
 CACAGCTGCTCTATTCTCACACAAATCTACCCCTGCGTGGCTGAACTGACGTTCCCTGGAGGTGTCCAGAAA
 GCTGATGTAACACAGGCCATAAAAGCTGCGTCTTAAGGCTGCCAGCGCTTGCCTAAAGGCTGCTGCTGCTGCTG
 AGAAGGCTCATGCCATTGACCCCTTAATTCTCTCTGTTGGCGAGCTGACAATGGCGAGGCTGAAGGCAAT
 GCAAGCTGCACAGTCAGTCTAGGGGTGCCAATATGGCAGAGAACCCACAAGCCATGATCTGCAACTCAATCCC
 AGTGAGAACTGCACCTGGACAATAGAAAGACCAAGAAAACAAAGCATTGAGAATTATCTTTCTATGTCCAGCTT
 GATCCAGATGGAAGCTGAAAGTAAAGCATTAAGCTTGTGACGGAACTCCAGCAATGGGCTCTGCTAGGG
 CAAGTCTGCAGTAAAAACGACTATGTTCTGTATTGAAATCATCATCCAGTACATTGACGTTCAAAATAGTTACT
 GACTCAGCAAGAATTCAAAGAACCTGTCTTGCTTCTACTACTCTCTCTCTAACATCTTCTATCCAAACTGT
 GGCGGTTACCTGGATACCTTGGAAAGGATCCTTACCCAGCCCCAATTACCCAAAGCCGATCTGAGCTGGCTTAT
 TGTGTGTGGCACATACAAGTGGAGAAAGATTACAAGATAAAACTAAACTCAAAAGAGATTTCCTAGAAATAGAC
 AAACAGTGCAAATTGATTTCTTGCATCTATGATGGCCCTCCACCAACTCTGGCTGATGGACAAGTCTGT
 GGCGTGTGACTCCCACCTCGAACGTCATCAAACCTCTGACTGTGCTGTTGCTACAGATTATGCCAAATTCT
 TACCGGGGATTCTGCTTCTACACCTCAATTATGAGAAAACATCAACACTACATCTTAACTTGCTCTCT
 GACAGGATGAGAGTTATTATAAGCAAATCTAACAGGCTTTAACTCTAAATGGAATAACTTGCAACTAAAA
 GACCCAACTTGCAGACAAAATTATCAAATGTTGGAATTCTGCTCCCTTAATGGATGGTACAATCAGA
 AAGGTAGAAGATCAGTCATTACTACACCAATAATCACCTTCTGACATCCTCAACTCTGAAGTGTACCC
 CGTCAGAACAACTCCAGATTATTGTGAAGTGTGAAATGGACATAATTCTACAGTGGAGATAATACATAACA
 GAAGATGATGTAATACAAGTCAAATGCACTGGCAAATAACACCCAGCATGGCTTTGAATCCAATTCA
 TTGAAAAGACTATACTTGAATCACCATATTATGTGGATTGAAACCAACTCTTTGTTCAAGTTAGTCTGCAC
 ACCTCAGATCCAAATTGGTGGTTCTGTACACTGTAGGCTCTCCACCTGACTTTGCATCTCCAACC
 TACGACCTAATCAAGAGTGGATGTAGTCGAGATGAAACTTGTGAAGGTGATCCCTTATTTGGACACTATGGGAGA
 TTCCAGTTAATGCCCTAAATTCTGAGAAGTATGAGCTCTGTGATCTGAGTGTAAAGTTGATATGTGAT
 AGCAGTGCACCAACAGTCTCGCTGCAATCAAGGTTGTGTCCTCCAGAAGCAAACGAGACATTCTTCAATATAATGG
 AAAACAGATTCCATCATAGGACCCATTGCTGAAAAGGGATCGAAGTGCAGTGGCAATTCAAGGATTTCAGCAT
 GAAACACATGCGGAAGAAACTCCAAACCAAGCCCTTCAACAGTGTGCATCTGTTCTCATGGTCTAGCTG
 AATGTGGTACTGTAGCGACAATCACAGTGAGGCATTGTAATCAACGGCAGACTACAAATACCAGAAGCTG
 CAGAACTATTAACTAACAGGTCCAACCCCTAACAGTGAGACATGTTCTCAGGATGCCAAAGGAAATGCTACCTCGT
 GGCTACACATATTGAAATAATGAGGAAGGGCCTGAAAGTGCACACACAGGCCATGTA

FIGURE 70

MELVRRILMPLTLLILSCLAEALTMAEAEGNASCTVSLGGANMAETHKAMILQLNPSENCTWTI
ERPENKSIRIIFSYVQLDPDGSCESENIKVFDGTSSNGPLLGQVCSKNDYVPVFESSSSTLT
FQIVTDSARIQRTVFVFYYFSPNISIPNCGGYLDTLEGSFTSPNPKPHPELAYCVWHIQV
EKDYKIKLNFKEIFLEIDKQCKFDLAIYDGPSTNSGLIGQVCGRVTPTFESSNSLTVVLS
TDYANSYRGFSASYTSIYAENINTTSLTCSSDRMRVIISKSYLEAFNSNGNNLQLKDPTCRP
KLSNVVEFSVPLNGCGTIRKVEDQSITYTNIITFSASSTSEVITRQKQLQIIVKCEMGHNST
VEIIYITEDDVIQSQNALGKYNTSMALFESNSFEKTILESPYYVDLNQTLFVQVSLHTSDPN
LVVFLDTCRASPTSDFASPTYDLIKSGCSRDETCKVYPLFGHYGRFQFNFKFLRSMSSVYL
QCKVLICDSSDHQSRCNQGCVSRSKRDISSYKWKTDSIIGPIRLKRDRSASGNSGFQHETHA
EETPNQPFNSVHLFSFMVLALNVVTVATITVRHFVNQRADYKYQKLQNY

Signal sequence:

amino acids 1-24

Transmembrane domain:

amino acids 571-586

N-glycosylation site.

amino acids 29-33, 57-61, 67-71, 148-152, 271-275, 370-374,
394-398, 419-423

Casein kinase II phosphorylation site.

amino acids 22-26, 108-112, 289-293, 348-352, 371-375, 379-383,
408-412, 463-467, 520-524, 556-560

Tyrosine kinase phosphorylation site.

amino acids 172-180, 407-415, 407-416, 519-528

N-myristoylation site.

amino acids 28-34, 38-44, 83-89, 95-101, 104-110, 226-232

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 7-18

FIGURE 71

GACGGAAGAACAGCGCTCCCGAGGCCGCGGGAGCCTGCAGAGAGGACAGCCGGCCTGCGCCG
 GGACATGCGGCCCGAGGAGCTCCCAGGCTCGCGTCCCGTGTGCTGCTGTTGCTGC
 TGCTGCCGCCGCCGTGCCCTGCCACAGCGCCACGCGCTTCGACCCCACCTGGGAGTCC
 CTGGACGCCGCCAGCTGCCCGGTGGTGTGACCAGGCCAAGTCGGCATCTTCATCCACTG
 GGGAGTGTTCGGTGTGCCCCAGCTCGTAGCGAGTGGTCTGGTGGTATTGGCAAAAGAAA
 AGATACCGAAGTATGTGGAATTATGAAAGATAATTACCCCTCTAGTTCAAATATGAAGAT
 TTTGGACCACTATTTACAGAAAATTTTAATGCCAACAGTGGCAGATATTTTCAGGC
 CTCTGGTGCCAAATACATTGCTTAACCTCCAAACATCATGAAGGCTTACCTGTGGGGT
 CAGAATATTCGTGGAACTGGAATGCCATAGATGAGGGGCCAAGAGGGACATTGTCAAGGAA
 CTTGAGGTAGCCATTAGGAACAGAACTGACCTGCGTTGGACTGTACTATTCCCTTTGA
 ATGGTTTCATCCGCTCTTCCTTGAGGATGAATCCAGTTCACTCCATAAGCGGCAATTCCAG
 TTTCTAAGACATTGCCAGAGCTCTATGAGTTAGTGAACAACATCAGCCTGAGGTTCTGTGG
 TCGGATGGTGACGGAGGACCCGGATCAAACTGGAACAGCACAGGCTTCTGGCCTGGTT
 ATATAATGAAAGCCCAGTTGGGGCACAGTAGTCACCAATGATCGTGGGGAGCTGGTAGCA
 TCTGTAAGCATGGTGGCTTCTATACCTGCAGTGTATACCCAGGACATCTTTGCCA
 CATAAATGGGAAAACATGCATGACAATAGACAAACTGTCCTGGGCTATAGGAGGGAAAGCTGG
 AATCTCTGACTATCTTACAATTGAAGAATTGGTAGCAAGCAACTTGTAGAGACAGTTCACTGTG
 GAGGAAATCTTTGATGAATATTGGGCCACACTAGATGGCACCATTCTGTAGTTTGAG
 GAGCGACTGAGGCAAGTGGGCTCTGGCTAAAGTCATGGAGAAGCTATTATGAAACCTA
 TACCTGGCGATCCAGAAATGACACTGTCACCCAGATGTGTGGTACACATCCAAGCCTAAAG
 AAAAATTAGTCTATGCCATTTCCTTAAATGGGCCACATCAGGACAGCTGTCCTGGCCAT
 CCCAAAGCTATTCTGGGGCAACAGAGGTGAAACTACTGGGCCATGGACAGCCACTTAAC
 GATTCTTGGAGCAAATGGCATTATGGTAGAACTGCCACAGCTAACATTCACTCAGATGC
 CGTGTAAATGGGCTGGCTTAGCCTAACTAATGTGATCTAAAGTGCAGCAGAGTGGCTG
 ATGCTGCAAGTTATGTCTAAGGCTAGGAACATCAGGTGTCTATAATTGTAGCACATGGAGA
 AAGCAATGTAAACTGGATAAGAAAATTATTGGCAGTTCCAGGCCCTTCCCTTTCCCACTA
 AATTTCCTTAAATTACCCATGTAACCATTAACTCTCCAGTGCACTTGCCATTAAAGTC
 TCTTCACATTGATTGTTCCATGTGACTCAGAGGTGAGAATTTCACATTATAGTAG
 CAAGGAATTGGTGGTATTATGGACCGAACTGAAAATTATGTTGAAGCCATATCCCCCATG
 ATTATATAGTTATGCATCACTTAATATGGGATATTCTGGAAATGCATTGCTAGTCAT
 TTTTTTGCCACATCATAGAGTGTATTACAAATCCTAGATGGCATAGCCTACTACA
 CACCTAATGTGTAGGTATAGACTGTTGCTCTGGCTACAGACATATACAGCATGTTACTG
 AATACTGTAGGCAATAGTAACAGTGGTATTGTATATGAAACATATGGAAACATAGAGAAG
 GTACAGTAAAATACTGTAAAATGGTCACCTGTATAGGGCACTTACCAAGAATGGAG
 CTTACAGGACTGGAAGTTGCTCTGGGTGAGTCAGTGAGTGAATGTAAGGCCTAGGACATTA
 TTGAACACTGCCAGACGTATAAATACTGTATGCTTAGGCTACACTACATTATAAAAAAAA
 GTTTTCTTCTTCATTATAAAATTACATAAGTGTACTGTAACTTACAAACGTTTAATT
 TTTAAACCTTTGGCTTTGTAATAACACTTAGCTAAAACATAAAACTCATTGTGCAA
 ATGTAA

FIGURE 72

MRPQELPRLAFPLLLLLLLLLPPPPCPAHSATRFDPTWESLDARQLPAWFDQAKFGIFIHWG
VFSVPSFGSEWFWWYQKEKIPKYVEFMKDNYPPSFKYEDFGPLFTAKFFNANQWADIQAS
GAKYIVLTSKHHEGFTLWGSEYSWNNAIDEGPKRDIVKELEVAIRNRTDLRGFLYYSLFEW
FHPLFLEDESSSFHKRQFPVSKTLPELYELVNNYQPEVLWSDGDGGAPDQYWNSTGFLAWLY
NESPVRGTVVTNDRGAGSICKHGGFYTCSDRYNPGHLLPHKWENCMTIDKLSWGYRREAGI
SDYLTIEELVKQLVETVSCGGNLLMNIGPTLDGTISVVFEERLRQVGWSLKVNGEAIYETY
WRSQNDTVTPDVWYTSKPKEKLVYAIFLKWPTSGQLFLGHPKAILGATEVKLLGHGQPLNWI
SLEQNGIMVELPQLTIHQMPCKWGWLALTNV

Signal sequence:

amino acids 1-28

N-glycosylation site.

amino acids 171-175, 239-243, 377-381

Casein kinase II phosphorylation site.

amino acids 32-36, 182-186, 209-213, 227-231, 276-280, 315-319,
375-375

Tyrosine kinase phosphorylation site.

amino acids 361-369, 389-397

N-myristoylation site.

amino acids 143-149, 178-184, 255-261, 272-278, 428-434

Leucine zipper pattern.

amino acids 410-432

Alpha-L-fucosidase putative active site.

amino acids 283-295

FIGURE 73

AGCAGGGAAATCCGGATGTCTCGGTTATGAAGTGGAGCAGTGAGTGTGAGCCTAACATAGT
 TCCAGAACTCTCCATCCGGACTAGTTATTGAGCATCTGCCTCTCATATCACCAGTGGCCATC
 TGAGGTGTTCCCTGGCTCTGAAGGGTAGGCACGATGGCCAGGTGCTTCAGCCTGGTGTG
 CTTCTCACCCATCTGGACACAGGGCTCTGGTCCAAGGCTCTTGCAGCAGAAGAGCT
 TTCCATCCAGGTGTCAAGAATTATGGGATCACCCCTGTGAGCAGAAGAGCAGC
 AGCTGAATTCACAGAAGCTAAGGAGGCCTGAGGCTGCTGGACTAAGTTGGCCGGCAAG
 GACCAAGTTGAAACAGCCTGAAAGCTAGCTTGAAACTGCAGCTATGGCTGGGTTGGAGA
 TGGATTGTGGTCATCTAGGATTAGCCAAACCCAAAGTGTGGAAAAATGGGTTGGTG
 TCCTGATTGGAAGGTTCCAGTGAGCCAGTTGCAGCCTATTGTTACAACATCTGAT
 ACTTGGACTAACTCGTCATTCCAGAAATTATCACCACCAAAGATCCCATATTCAACACTCA
 AACTGCAACACAAACAGAATTATTGTCAAGTACAGTACCTACTCGGTGGCATCCCTT
 ACTCTACAATACCTGCCCTACTACTACTCCTCCTGCTCCAGCTTCACTTCTATTCCACGG
 AGAAAAAAATTGATTGTGTCAGAGTTTATGAAACTAGCACCAGTCTACAGAAC
 TGAACCATTTGTTGAAAATAAAGCAGATTCAAGAATGAAGCTGCTGGGTTGGAGGTGTCC
 CCACGGCTCTGCTAGTGCTCTCCCTTGGTGTGAGCTGGTCTGGATTTC
 TATGTCAAAAGGTATGTGAAGGCCTCCCTTACAAACAAGAATCAGCAGAAGGAAATGAT
 CGAAACCAAAGTAGTAAAGGAGGAAGGCAATGATAGAACCTTAATGAGGAATCAAAGA
 AACTGATAAAAACCCAGAAGAGTCCAAGAGTCCAAGCAAAACTACCGTGCATGCCTGGAA
 GCTGAAGTTTAGATGAGACAGAAATGAGGAGACACACCTGAGGCTGGTTCTTCTGCTCC
 TTACCCCTGCCAGCTGGGAAATCAAAGGGCAAAGAACCAAAGAACAGTCCACCCCTT
 GGTTCTAACTGGAATCAGCTCAGGACTGCCATTGACTATGGAGTGCACCAAAGAGAAC
 CCTTCTCCTTATTGTAACCCCTGTCTGGATCCTATCCTCACCTCCAAAGCTCCCACGGCC
 TTTCTAGCTGGCTATGCTTAATAATATCCACTGGAGAAAGGAGTTTGCAAAAGTCAA
 GGACCTAAAACATCTCATCAGTATCCAGTGGTAAAAGGCCCTCTGGCTGTGAGGCTAGG
 TGGTTGAAAGCCAAGGAGTCAGGACCAAGGCTTCTACTGATTCCGCAGCTCAGAC
 CCTTCTCAGCTGTGAAAGAGAACACGTATCCACCTGACATGCTCTGAGCCGGTA
 AGAGCAAAAGAATGGCAGAAAAGTTAGCCCTGAAAGCCATGGAGATTCTCATAACTGAG
 ACCTAATCTCTGAAAGCTAAAATAAGAAATAGAACACAGGCTGAGGATACGACAGTACACT
 GTCAGCAGGGACTGTAAACACAGACAGGGTCAAAGTGTCTCTGAACACATTGAGTTGGA
 ATCACTGTTAGAACACACACACTTACTTTCTGGTCTCTACCACTGCTGATATTTCTCT
 AGGAAATATACTTTACAAGTAACAAAAATAAAACTCTTATAAAATTCTATTTTATCTGA
 GTTACAGAAATGATTACTAAGGAAGATTACTCAGTAATTGTTAAAAGTAATAAAATTCA
 ACAAAACATTGCTGAATAGCTACTATATGTCAAGTGTGCAAGGTATTACACTCTGTAAT
 TGAATATTATTCTCAAAAAATTGCACATAGTAGAACGCTATCTGGGAAGCTATTTTCA
 GTTTGATATTCTAGCTTATCTACTTCCAAACTAATTGTTATTGCTGAGACTAATCTT
 ATTCACTTCTCTAATATGCCAACATTATAACCTTAATTATTATAACACACCTAAGAAG
 TACATTGTTACCTCTATACCAAAAGCACATTAAAAGTGCCTTAACAAATGTATCACTA
 GCCCTCCTTTCCAACAAGAAGGACTGAGAGATGCAGAAATATTGTGACAAAAATTAA
 AGCATTAGAAAACCTT

FIGURE 74

MARCFSLVLLTSIWTTRLLVQ GSLRAEELSIQVSCRIMGITLVS K KANQQLNFTEAKBACR
LLGLSLAGKDQVETALKASFETCSYGVGDGFVVISRISP NPKCGKNGVGVLIW KVPVSRQF
AAYCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSV ASPYSTI PAPTTPP
APASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKA AFKNEAAGFGGVPTALLVLALLFF
GAAAGLGFCYVKRYVKAFPFTNKNQQKEMIETKVVKEEKANDSNPNEESKKTDKNPEESKSP
SKTTVRCLEAEV

Signal sequence:

amino acids 1-16

Transmembrane domain:

amino acids 235-254

N-glycosylation site.

amino acids 53-57, 130-134, 289-293

Casein kinase II phosphorylation site.

amino acids 145-149, 214-218

Tyrosine kinase phosphorylation site.

amino acids 79-88

N-myristoylation site.

amino acids 23-29, 65-71, 234-240, 235-239, 249-255, 253-259

FIGURE 75

AGATGGCGGTCTTGGCACCTCTAATTGCTCTCGTGTATTCGGTGCCCGACTTCACGATGG
CTCGCCCAACCTTACTACCTTCTGTCGGCCCTGCTCTGCTGCCTTCCTACTCGTGAGGAA
ACTGCCGCCGCTCTGCCACGGCTGCCACCCAACCGCGAAGACGGTAACCGTGTGACTTTG
ACTGGAGAGAAGTGGAGATCCTGATGTTCTCAGTGCATTGTGATGATGAAGAACCGCAGA
TCCATCACTGTGGAGCAACATATAGGCAACATTTCATGTTAGTAAAGTGGCCAACACAAT
TCTTTCTTCGCTTGGATATTGCATGGGCCTACTTACATCACACTCTGCATAGTGTCC
TGATGACGTGCAAACCCCCCTATATGGGCCTGAGTATATCAAGTACTTCAATGATAAA
ACCATTGATGAGGAACTAGAACGGACAAGAGGGTCATTGGATTGTGGAGTTCTTGCCAA
TTGGTCTAATGACTGCCAATCATTGCCCTATCTATGCTGACCTCTCCCTAAATACAAC
GTACAGGGCTAAATTTGGGAAGGTGGATGTTGGACGCTACTGATGTTAGTACGGTAC
AAAGTGAGCACATCACCCCTACCAAGCAACTCCCTACCCCTGATCCTGTTCAAGGTGGCAA
GGAGGCAATGCGGCGGCCACAGATTGACAAGAAAGGACGGGCTGTCTCATGGACCTCTG
AGGAGAATGTGATCCGAGAATTAACTTAAATGAGCTATACCAGCGGCCAAGAAACTATCA
AAGGCTGGAGACAATATCCCTGAGGAGCAGCCTGTTCAACCCCCACCAAGTGTCA
TGGGGAAAACAAGAAGGATAAATAAGATCCTCACTTGGCAGTGCTCCTCTCCTGTCAATT
CCAGGCTTTCCATAACCACAAGCCTGAGGCTGCAGCCTTNATTNATGTTTCCCTTGG
CTGNGACTGGNTGGGCAGCATGCAGCTCTGATTTAAAGAGGCATCTAGGAAATTGTCAG
GCACCCCTACAGGAAGGCCTGCCATGCTGTGGCCAAGCTGTTCACTGGAGCAAGAAAGAGATC
TCATAGGACGGAGGGGGAAATGGTTCCCTCCAAGCTTGGTCAGTGTGTTACTGCTTATC
AGCTATTCAAGACATCTCCATGGTTCTCCATGAAACTCTGTGGTTCATCATTCTCTTAG
TTGACCTGCACAGCTTGGTAGACCTAGATTAAACCTAAGGTAAAGATGCTGGGTATAGAA
CGCTAAGAATTTCACCCCAAGGACTCTGCTTCCATTAGCCCTCTGGCTTATGGTC
TTCATTAAGTATAAGCCTAACTTGTGCTAGTCCTAAGGAGAAACCTTAACCACAAAG
TTTTTATCATTGAAGACAATATTGAACAACCCCTATTTGTGGGATTGAGAAGGGTGAA
TAGAGGCTTGAGACTTCCCTTGTGTGGTAGGACTGGAGGAGAAATCCCTGGACTTCAC
TAACCCCTGACATACTCCCCACACCCAGTTGATGGCTTCCGTAATAAAAGATTGGGATT
TCCTTTG

FIGURE 76

MAVLAPLIALVYSPVRLSRWLAQPYYLLSALLSAAFLLVRKLPPCHGLPTQREDGNPCDFD
WREVEILMFLSAIVMMKNRRSITVEQHIGNIFMFSKVANTILFFRLDIRMGLLYITLCIVFL
MTCKPPLYMGPEYIKYFNDKTIDEELERDKRVTWIEFFANWSNDCQSFAPIYADLSLKYNC
TGLNFGKVDVGRYTDVSTRYKVSTSPLTKQLPTLILFQGGKEAMRRPQIDKKGRAVSWTFSE
ENVIREFNLNELYQRAKKLSKAGDNIPEEQPVASTPTVSDGENKKDK

Signal sequence:

amino acids 1-48

Transmembrane domain:

amino acids 111-125

N-glycosylation site.

amino acids 165-169, 185-189

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 154-158, 265-269

Casein kinase II phosphorylation site.

amino acids 51-55, 145-149, 245-249, 286-290, 288-292

N-myristoylation site.

amino acids 188-194, 225-231

Myb DNA-binding domain repeat signature 1.

amino acids 244-253

FIGURE 77

GGACAGCTCGGGCCCCGAGAGCTAGCCGTGAGGAGCTGCCTGGGACGTTGCCCTG
 GGGCCCCAGCCTGGCCGGTCACCCCTGGCATGAGGAGATGGGCCTGTTGCTCCTGGTCCCA
 TTGCTCCTGCTGCCCGCTCCTACGGACTGCCCTCTACAACGGCTCTACTACTCCAACAG
 CGCCAACGACCAGAACCTAGGCAACGGCATGGCAAAGACCTCCTTAATGGAGTGAAGCTGG
 TGGTGGAGACACCCGAGGAGACCCCTGTTCACCAAGGGGCCAGTGTGATCCTGCCCTGC
 CGCTACCGCTACGAGCCGGCCCTGGTCTCCCCCGGGCGTGTGCGTGTCAAATGGTGGAAAGCT
 GTCGGAGAACGGGCCCCAGAGAAGGACGTGCTGGTGGCATGGGCTGAGGCACCGCTCCT
 TTGGGACTACCAAGGCCCGTGCACCTGCCGAGGACAAAGAGCATGACGTCTCGCTGGAG
 ATCCAGGATCTGCCGCTGGAGGACTATGGCGTTACCGCTGTGAGGTCAATTGACGGCTGGA
 GGATGAAAGCGGTCTGGTGGAGCTGGAGCTGCCGGGTGTGGCTTTCTTACCAAGCTCCCA
 ACGGGCGTACCAAGTTCAACTCCACGAGGGCCAGCAGGTCTGTGCAGAGCAGGCTGCCGTG
 GTGGCCTCTTGAGCAGCTCTCCGGGCTGGAGGGAGGGCCTGGACTGGTGCAACGCCGG
 CTGGCTGCAGGATGCTACGGTCAGTACCCCATCATGTTGCCCGCAGCCCTGCCGTGGCC
 CAGGCCTGGCACCTGGCGTGCAGCTACGGCCCCGCCACGCCGCTGCACCGCTATGAT
 GTATTCTGCTCGCTACTGCCCTCAAGGGCGGGTGTACTACCTGGAGCACCCGTGAGAACCT
 GACGCTGACAGAGGAAGGGAGGCCTGCCAGGAAGATGATGCCACGATGCCAAGGTGGAC
 AGCTTTGCCGCTGGAAGTTCCATGGCCTGGACCGCTGCACGCTGGCTGGCAGAG
 GGCAGCGTCCGCTACCCCTGGTTACCCGCATCCTAACTGTGGGCCCCAGAGCCTGGGTT
 CCGAAGCTTGCTTCCCCGACCCGCAGAGCCCTGTACGGTGTACTGCTACCGCCAGC
ACTAGGACCTGGGCCCCCTCCCTGCCGATTCCCTCACTGGCTGTGTTAGTATTGAGTGGTT
 CGTTTCCCTTGTTGGAGCCATTAACTGTTTATACTCTCAATTAAATTTCT
 TTAAACATTTTTACTATTTTGAAAGCAAACAGAACCCATGCCCTCCCTTGCTCCTG
 GATGCCCAACTCCAGGAATCATGCTTGCCTCCCTGGCCATTGCGGTTTGCGGCTCTG
 GAGGGTCCCCGCCATCCAGGCTGGCTCCCTCCCTTAAGGAGGTTGGTGCCAGAGTGGC
 GGTGGCCTGTCTAGAATGCCGCCGGAGTCGGGCATGGTGGCACAGTCTCCCTGCC
 CAGCCTGGGGAAGAAGAGGGCCTGGGGCCTCCGGAGCTGGCTTGGCCTCTCCTGCC
 CACCTCTACTTCTGTGAAGCCGCTGACCCAGTCTGCCACTGAGGGCTAGGGCTGGAA
 GCCAGTTCTAGGCTTCCAGGCAGAATCTGAGGAGGAAGAAACTCCCTCCCCGTTCCCT
 TCCCTCTGGTCCAAAGAATCTGTTGTCATTGTTCTCCTGTTCCCTGTGTTG
 GGAGGGCCCTCAGGTGTGTACTTGGACAATAATGGTGCTATGACTGCCCTCCGCCAA
 AAA
 AAA

FIGURE 78

MGLLLLVPLLLLPGSYGLPFYNGFYYSNSANDQNLGNHGKDLLNGVKLVVETPEETLFTYQ
GASVILPCRYRYEPALVSPRRVRVKWWKLSENGAPEKDVLVAIGLRHRSFGDYQGRVHLRQD
KEHDVSLEIQDLRLEDYGRYCEVIDGLEDESGLVELELRGVVFVFPYQSPNGRYQFNFHEGQQ
VCAEQAAVVASFEQLFRAWEGLDWCNAGWLQDATVQYPIMLPRQPCGGPGLAPGVRSYGPR
HRRLHRYDVFCFATALKGRVYYLEHPEKLTILTEAREACQEDDATIAKVGQLFAAWKFHGLDR
CDAGWLADGSVRYPVVHPHPNCGPPEPGVRSFGFPDPQSRLYGVYCYRQH

Signal sequence:

amino acids 1-17

Casein kinase II phosphorylation site.

amino acids 29-33, 53-57, 111-115, 278-282

Tyrosine kinase phosphorylation site.

amino acids 137-145

N-myristoylation site.

amino acids 36-42, 184-190, 208-214, 237-243, 297-303, 307-313

FIGURE 79

FIGURE 80

MMWRPSVLLLLLRLHGAQGKPSPDAGPHQGRVHQAPLSDAPHDDAHGNFQYDHEAFLGR
EVAKEFDQLTPEESQARLGRIVDRMDRAGDGWVSLAELRAWIAHTQQRHIRDSVSAAWDT
YDTDRDGRVGWEELRNATYGHYAPGEEFHDVEDAETYKKMLARDERRFRVADQDGDSMATRE
ELTAFLHPEEFPHMRDIVAETLEDLDRNKGYVQVEEYIADLYSAEPGEEEPAWVQTERQQ
FRDFRDLNKGHDGSEVGHVLPPAQDQPLVEANHLLHESDTDKDGRLSKAEILGNWNMFV
GSQATNYGEDLTRHHDEL

Signal sequence:

amino acids 1-20

N-glycosylation site.

amino acids 140-144

Casein kinase II phosphorylation site.

amino acids 72-76, 98-102, 127-131, 184-188, 208-212, 289-293,
291-295, 298-302

N-myristoylation site.

amino acids 263-269, 311-317

Endoplasmic reticulum targeting sequence.

amino acids 325-330

FIGURE 81

GGGGCCTTGCCTCCGCACTCGGGCGCAGCCGGTGGATCTGAGCAGGTGCGGAGCCCCGG
 GCGGGCGGGCGCGGGTGCAGGGATCCCTGACGCCCTGTCCCTGTTCTTGTGCTCCAG
 CCTGTCGTCGTCGTTGGCGCCCCCGCTCCCGCGGTGCGGGGTTGCACACCGATCCTG
 GGCTTCGCTCGATTGCCGCCAGGCCCTCCAGACCTAGAGGGCGCTGGCCTGGAGCAG
 CGGGTCGCTGTGTCCTCTCCTCTGCCGCCGCCCCGGGATCCGAAGGGTGCAGGGCTCT
 GAGGAGGTGACGCCGGGGCCTCCGCACCCCTGCCCTGCCGCATTCTCCCTCTCCAG
 GTGTGAGCAGCCTATCAGTCACCATGTCCGCAGCCGGATCCCGCTCTGCCCTGGTGTG
 TGTCTGCTGCTGCCGGGGCCCGGGCAGCGAGGGAGCCGCTCCATTGCTATCACATG
 TTTTACCAAGAGGCTTGGACATCAGGAAAGAGAAAGCAGATGTCCTCTGCCAGGGGCTGCC
 CTCTGAGGAATTCTCTGTATGGGAACATAGTATATGCTCTGTATCGAGCATATGTGGG
 GCTGCTGCCACAGGGAGTAATCAGCAACTCAGGGGACCTGTACAGAGTCTATAGCCTACC
 TGGTCGAGAAAATATTCCCTAGTAGATGCCATGGCATCCAGTCTCAAATGCTTCTAGAT
 GGTCTGCTTCTTCACAGTAACAAAGGCAAAAGTAGTACACAGGAGGCCACAGGACAAGCA
 GTGTCCACAGCACATCCACCAACAGTAAACGACTAAAGAAAACACCCGAGAAGAAAATGG
 CAATAAAAGATTGTAAGCAGACATTGCATTTCTGATTGATGGAAGCTTAAATTGGGCAGC
 GCCGATTAAATTACAGAAGAATTGGGAAAGTGGCTTAATGTTGGGAAATTGGAAACA
 GAAGGACCACATGTGGGCCTGTTCAAGCCAGTGAACATCCAAAATAGAATTACTTGAA
 AAACATTACATCAGCCAAAGATGTTGTTGCCATAAAGGAAGTAGGTTTCAGAGGGGTA
 ATTCCAATACAGGAAAAGCCTTGAAGCATACTGCTCAGAAATTCTCACGGTAGATGCTGGA
 GTAAGAAAAGGGATCCCCAAAGTGGTGGTATTATTGATGGTGGCCTCTGATGACAT
 CGAGGAAGCAGGCATTGTGGCCAGAGAGTTGGTCAATGTATTATAGTTCTGTGGCCA
 AGCCTATCCCTGAAGAACTGGGATGGTCAGGATGTCACATTGTTGACAAGGCTGTCTGT
 CGGAATAATGGCTTCTCTTACACATGCCCAACTGGTTGGCACCAAAATACGTAAA
 GCCTCTGGTACAGAAGCTGTGCACTCATGAACAAATGATGTGCAAGCAAGACCTGTTATAACT
 CAGTGAACATTGCCCTCTAATTGATGGCTCCAGCAGTGGAGATAGCAATTCCGCCTC
 ATGCTTGAATTGTTCCAACATAGCCAAGACTTTGAAATCTGGACATTGGTGCAGAT
 AGCTGCTGTACAGTTACTTATGATCAGCGCACGGAGTTCAAGTTCACTGACTATAGCACCA
 AAGAGAAATGTCCTAGCTGTATCAGAAACATCCGCTATATGAGTGGGAACAGCTACTGGT
 GATGCCATTCTTCACTGTTAGAAATGTGTTGGCCCTATAAGGGAGAGCCCCAACAGAA
 CTTCTAGTAATTGTCACAGATGGCAGTCCTATGATGATGTCAGGCTCTGAGCTGCTG
 CACATGATGCAGGAATCACTATCTCTGTTGGTGTGGCTGGCACCTCTGGATGACCTG
 AAAGATATGGCTCTAAACCGAAGGAGTCTCACGCTTCTCACAAGAGAGTTCACAGGATT
 AGAACCAATTGTTCTGATGTCATCAGAGGCATTGTTAGAGATTTCTAGAATCCCAGCAAT
AAATGGTAACATTGACAACAGAAAGAAAAGTACAAGGGATCCAGTGTAAATTGTATT
 CTCATAACTGAAATGCTTCACTGATAGAATCAGATAACAAACTATTAAGTATGTCAAC
 AGCCATTAGGCAAATAAGCACTCCTTAAAGCCGCTGCCCTCTGGTACAATTACAGTGT
 ACTTTGTTAAAACACTGCTGAGGCTTCATAATCATGGCTCTAGAAAATCAGGAAAGAGGA
 GATAATGTGGATTAAAACCTAAGAGTTCAACCATGCCTACTAAATGTACAGATATGCAAA
 TTCCATAGCTAATAAAAGAATCTGATACTTAGACCAAAAAAA

FIGURE 82

MSAAWIPALGLGVCLLLLPGPAGSEAAPIAITCFTRGLDIRKEKADVLCPGGCPLLEFSVY
GNIVYASVSSICGAAVHRGVISNSGGPVRYSLPGRENYSSVDANGIQSQMLSRWSASFTVT
KGKSSTQEATGQAVSTAHPPTGKRLKKTPEKKTGNKDCDKADIAFLIDGSFNIGQRRFNLQKN
FVGKVALMLGIGTEGPHVGLVQASEHPKIEFYLKNFTSAKDVLFAIKEVGFRGGNSNTGKAL
KHTAQKFFTVDAGVRKGIPKVVVFIDGWPSSDDIEAGIVAREFGVNVFIVSVAKPIPEELG
MVQDVTFVDKAVCRNNNGFFSYHMPNWFGTTKYVKPLVQKLCTHEQMMCSKTCYNSVNIACLI
DGSSSVGDSNFRMLFEVSNIAKTFEISDIGAKIAAVQFTYDQRTEFSFTDYSTKENVLAVI
RNIRYMSGGTATGDAISFTVRNVFGPIRESPNKNFLVIVTDGQSYDDVQGPAAAHDAGITI
FSVGVAWAPLDDLKDMASKPKESHAFFTREFTGLEPIVSDVIRGICRDFLESQQ

Signal sequence:

amino acids 1-24

N-glycosylation site.

amino acids 100-104, 221-225

Casein kinase II phosphorylation site.

amino acids 102-106, 129-133, 224-228, 316-320, 377-381, 420-424,
425-429, 478-482, 528-532

N-myristoylation site.

amino acids 10-16, 23-29, 81-87, 135-141, 158-164, 205-211,
239-245, 240-246, 261-267, 403-409, 442-448, 443-449

Amidation site.

amino acids 145-149

FIGURE 83

CGCCGCGCTCCCGCACCGCGGGCCCGCCACCGCGCCGCTCCGCATCTGCACCCGAGCCC
 GGCAGGCTCCGGCGGGAGCGAGCAGATCCAGTCCGGCCCGCAGCGCAACTCGGTCCAGTCG
 GGGCGCGGCTGCGGGCGCAGAGCGAGATGCAGCGCTGGGGCACCCCTGCTGTGCCTGC
 TGCTGGCGCGGGTCCCCACGGCCCCCGCGCCGCTCCGACGGCGACCTCGGCTCCAGTC
 AAGCCCCGGCTCTCAGCTACCCGAGGAGGAGGACCCCTCAATGAGATGTTCCGCAG
 GGTTGAGGAACGTGATGGAGGACACGCAGCACAAATTGCGCAGCGGGTGGAAAGAGATGGAGG
 CAGAAGAAGCTGCTAAAGCATCATCAGAAGTGAACCTGGCAAACCTACCTCCAGCTAT
 CACAATGAGACCAACACAGACAGAAGGTTGAAATAATACCATCCATGTGCACCGAGAAAT
 TCACAAGATAACCAACAACAGACTGGACAATGGTCTTTCAGAGACAGTTACATCTG
 TGGGAGACGAAGAAGGCAGAAGGAGCCACGAGTGCATCATCGACGAGGACTGTGGGCCAGC
 ATGTAAGTGCAGTTGCCAGCTTCCAGTACACCTGCCAGCCATGCCAGGGCCAGAGGATGCT
 CTGCACCCGGGACAGTGAGTGCTGTGGAGACCAGCTGTGTCTGGGTCACTGCACCAAAA
 TGGCCACCAGGGGAGCAATGGGACCATCTGTGACAACCAGAGGGACTGCCAGCCGGGCTG
 TGCTGTGCCTTCCAGAGAGGGCTGCTGTTCCCTGTGTGCACACCCCTGCCGTGGAGGGGA
 GCTTGCACATGACCCGCCAGCCGGCTCTGGACCTCATCACCTGGAGCTAGAGCCTGATG
 GAGCCTGGACCGATGCCCTGTGCCAGTGGCCTCTGCCAGCCCCACAGCCACAGCCTG
 GTGTATGTGTGCAAGCCGACCTTCGTGGGGAGCCGTGACCAAGATGGGAGATCCTGCTGCC
 CAGAGAGGTCCCCGATGAGTATGAAGTGGCAGCTTCATGGAGGAGGTGCGCCAGGAGCTGG
 AGGACCTGGAGAGGAGCCTGACTGAAGAGATGGCCTGGGGAGCCTGCCAGCCGCT
 GCACTGCTGGAGGGGAAGAGATTAGATCTGGACCAGGCTGTGGGTAGATGTCAATAGAA
 ATAGCTAATTATTCAGGTGTGCTTACAGGCTGGCTGACAGCATGAGGTGTTGTGCAATTGTT
 TCTCTCCCAGTAAGTTCCTCTGGCTGACAGCATGAGGTGTTGTGCAATTGTT
 TCCCCCAGGCTGTTCTCCAGGCTTCACAGTCTGGCCTGGGAGAGTCAGGAGGGTTAAC
 TGCAGGAGCAGTTGCCACCCCTGTCCAGATTATTGGCTGCTTGCTCTACCAAGTGGCAG
 ACAGCCGTTGTTCTACATGGCTTGATAATTGTTGAGGGAGGAGATGGAAACAATGTGG
 AGTCTCCCTCTGATTGGTTGGGAAATGTGGAGAAGAGTGCCTGCTTGCAAAACATCAA
 CCTGGCAAAATGCAACAAATGAATTTCACAGCAGTTCTTCATGGCATAGGTAAGCTG
 TGCCCTCAGCTGTTGAGATGAAATGTTCTGTCACCCCTGCATTACATGTGTTATTCA
 AGCAGTGGCTCAGCTCACCTCTGGCCAGGGCAGCATTTCATATCCAAGATCAATT
 CCTCTCTCAGCACAGCCTGGGGAGGGGTATTGTTCTCGTCCATCAGGGATCTCAGAG
 GCTCAGAGACTGCAAGCTGCTGCCAAGTCACACAGCTAGTGAAGACCAGAGCAGTTCA
 CTGGTTGTGACTCTAAGCTCAGTGCTCTCCACTACCCACACCAGCCTGGTGCACCAA
 AAGTGTCCCCAAAAGGAAGGAGAATGGGATTTTCTTGAGGCATGCACATCTGAATTAAAG
 GTCAAACATAATTCTCACATCCCTCTAAAGTAAACTACTGTTAGGAACAGCAGTGTCTC
 AGTGTGGGCAGCCGCTCTAATGAAGACAATGATATTGACACTGTCCTCTGGCAGT
 TGCATTAGTAACCTTGAAAGGTATATGACTGAGCGTAGCAGCTAACCTGCAGAAACA
 GTACTTAGTAATTGTAGGGCGAGGATTATAAATGAAATTGCAAAATCACTAGCAGCAAC
 TGAAGACAATTATCAACCAACGTGGAGAAAATCAAACCGAGCAGGGCTGTGAAACATGGTT
 GTAATATGCGACTGCGAACACTGAACCTACGCCACTCCACAAATGATGTTTCAGGTGTCA
 TGGACTGTTGCCACCATGTTACATCCAGAGTCTTAAAGTTAAAGTGTGACATGATTGTA
 TAAGCATGCTTCTTGAGTTAAATTATGATAAACATAAGTGCATTAGAAATCAAGC
 ATAAATCACTGCAAAAAAAAAAAAAAA

FIGURE 84

MQRLGATLLCLLLAAAVPTAPAPAPTATSAPVKPGPALSYPQEEATLNEMFREVEELMEDTQ
HKLRSAVEEEMEAEEAAKASSEVNLPPSYHNETNTDTKVGNNNTIHVHREIHKITNNQTG
QMVFSETVITSVGDEEGRRSHECIIDEDCGPSMYCQFASFQYTCQPCRQMLCTRSECCG
DQLCVWGHCTKMATRGNSGTICDNQRDCQPGCCAFQRGLLFPVCTPLPVEGELCHDPASRL
LDLITWELEPDGALDRCPASGLLCQPHSHSLVYVCKPTFVGSRDQDGEILLPREVPDEYEV
GSFMEEVRQELEDLERSLTEEMALGEAAAAALLGEEI

Signal sequence:

amino acids 1-19

N-glycosylation site.

amino acids 96-100, 106-110, 121-125, 204-208

Casein kinase II phosphorylation site.

amino acids 46-50, 67-71, 98-102, 135-139, 206-210, 312-316,
327-331

N-myristoylation site.

amino acids 202-208, 217-223

Amidation site.

amino acids 140-144

FIGURE 85

AAGGAGGCTGGGAGGAAAGAGGTAAAGAAAGGTTAGAGAACCTACCTCACATCTCTGGGCTCAGAAGGACTCTA
AAGATAACAATAATTTCAGCCCCATCCACTCTCCTCCCTCCAAACACACATGTGATGTACACACACACATAC
CACACATACACCTTCTCTCCTCACTGAAGACTCACAGTCACTCACTCTGTGAGCAGGTATAGAAAAGGACAC
TAAAGCCTTAAGGACAGGCTGCCATTACCTCTGAGCTCCTTGGCTTGTGAGTCAAAAAACATGGGAGGGG
CCAGGCACGGTACTCACACCTGTAATCCAGCATTGGAGACCGAGGTGAGCAGATCACTTGAGGTAGGAG
TTCGAGACCAGCCTGCCAACATGGAGAAACCCCCATCTACTAAAAATACAAAATTAGCCAGGAGGTGGTGGC
AGGTGCCCTGTAATCCCAGTCACTCAGGTGAGCAGGAGAATCGCTTGAATCCAGGAGGCCAGGTGAGTGCAGT
CAGCTGAGTGCACCGCTGCACTCCAGGCTGGGTGACAGAATGAGACTCTGTCTAAACAAACACACGGGAGGA
GGGGTAGATACTGCTTCTGCAACCTCTTAACCTGCTCATCTCTTCCAGGGCTGCCCTGATGGGGCTG
GCAATGACTGAGCAGGCCAGCCCCAGAGGAACAGGAAGAGAAGGCATATTGAGGAGGGCAAGAAGTGACGCCCG
GTGAGAATGACTGCCCTGGGAGGGTGGTCTTGGGCCCTGGCAGGGTGTGCTGACCCCTTACCCGTCAAAACACA
AAGAGCAGGACTCCAGACTCTCTTGTGAATGGTCCCCGCCCCCTGAGCTCCACCATGAGGCTTCTGTGGCCCC
ACTCTTGCTAGCTGGGTGGCTGGTGCCTGCACTGCACTGTGCCCCGGTACCCCTGGCATGTTCCCTGCCCTCA
GTGTGCCCTGCCAGATCCGGCCCTGGTATACGCCCCGCTGCTCTACCGCAGGCTACCAACTGTGGACTGCAATGA
CCTATTCCGTACGGCAGTCCCCCGGCACTCCCCGAGGACACAGACCCCTGCTCTGAGCAGAACAGCATTGT
CCGTGTGGACCAGAGTGAAGCTGGTACCTGGCAATCTCACAGAGCTGGACCTGTCAGAAGCTTTCGG
TGCCCCGAGACTGTGATTTCATGCCCTGCCAGCTGCTGAGCCTGACCTAGAGGAGAACAGCTGACCCGGCT
GGAGGACCACAGCTTGCAAGGGCTGCCAGGCTACAGGAACTCTATCTCAACCACAAACCAGCTTACCGCAGC
CCCCAGGGCTTTCTGGCCTCAGCAACTTGCTGCCGTGCACCTCAACTCCAACCTCTGAGGGCATTGACAG
CCGCTGGTTGAAATGCTGCCAACCTGGAGATACTCATGATTGGGGCAACAAGGTAGATGCCATCTGGACAT
GAACCTCCGGCCCTGGCCAACCTGCGTAGGCTGGTGTAGCAGGATGAACCTGCGGGAGATCTCCGACTATGC
CTGGAGGGGCTGCAAAGCCTGGAGAGGCTCTCTTCTATGACAACCAAGCTGGCCGGGTGCCAGGGGCACT
GGAACAGGTGCCGGCTCAAGTCTTAGACCTCAACAAGAACCCGCTCCAGGGTAGGGCCGGGGACTTGC
CAACATGCTGCACCTTAAGGAGCTGGACTGAACAAACATGGAGGAGCTGGTCTCCATGACAAGTTGGCCCTGGT
GAACCTCCCCGAGCTGACCAAGCTGGACATCACAATAACCCACGGCTGTCTTCACTCCACCCCGCGCCCTCA
CCACCTGCCCTAGATGGAGACCCCTATGCTCAACAACAAACGCTCTCAGTGCCTGACCAAGCAGCGTGGAGTC
CTGGCCCAACCTGCAAGGAGGTAGGTCTCCACGGCAACCCATCCGCTGTACTGTGTATCCGCTGGGCAATGC
CACGGGCAACCGTGTCCGTTCATGAGCGCAATCCACCCGTGTGCGGAGCTCCGGACCTCCAGCGCCTCCC
GGTCCGTAGGTGCCCTCCGGAGATGACGGACACTGTTGCCCTCATCTCCCCACGAAGCTTCCCCCAAG
CCTCCAGGTAGCCAGTGGAGAGAGCATGGTGTGCAATTGCCGGACTGGCGAACCCGAACCGAGATCTACTG
GGTCACTCCAGCTGGCTTCAGTACACCTGCCATGCAAGCAGGAGGTACCCGGTGTACCCCGAGGGGACCC
GGAGCTGCGAGGGTGAACAGCAGAACAGGGCAGGGTATACACCTGTGTGGCCAGAACCTGGTGGGGCTGACAC
TAAGACGTTAGTGTGGTTGTGGGCCGTCTCCTCCAGCCAGGAGGGACGAAGGACAGGGGCTGGAGCTCCG
GGTGCAGGAGACCCACCCCTATCACATCTGCTATCTGGGTCAACCCACCCACAGTGTCCACCAACCTCAC
CTGGTCCAGTGCCTCTCCCTCCGGGCCAGGGGCCACAGCTCTGGCCCTGCCCTGGGAAACCCACAGCTA
CAACATTACCGCCTCTTCAGGCCACGGAGTACTGGGCCGTCTCAGGACACTTCTGGCCACAGGCTTAGGGGATGT
GTTGGCTTGTGATGGGCCAGGACCAAGAGGCCACTTCTGGCCAGGAGGTAGGGCCACCTTGGCACAGGCCA
TGCCATCTGGCTCGCTGTCTTCTCCAGCTGGCAGCTGGCTAGGGCCACCTTGGCACAGGCCAACCCAGGAA
GGGTGTGGGTGGGAGGCCGGCTCTCCCTCCAGCTGGCTTCTGGGCTGGAGTGGCCACCTTGTGCTGGGGTGT
GTCTGCTCCCTCGCTGTGCCCTGGAAATCCAGGGAGGAAGCTGCCAGATCCTCAGAAGGGGAGACACTGTTGCC
ACCATTGTCTCAAAATTCTTGAAGCTCAGCTGCTTCTCAGCAGTAGAGAAATCACTAGGACTACTTTTACCAA
AGAGAACAGTCTGGCCAGATGCCCTGCCAGGAAGGGACATGGACCCACGTGCTTGGGCCAGCTGGCAGCTGGC
CAAGACAGATGGGCTTGTGGCCCTGGGGTCTGCAAGCTTGAAGGAAACATCTCCAGGAGGACTTGGCTAG
CCTCTGCTGCCATTCTGAGGAACATCTCAAGGAACAGGAGGGACTTGGCTAGAGCCTCTGCCCTCCATCT
CTCTCTGCCAGAGGCTCTGGCCCTGGCTGGCTGGCTTCTGGCTACCTGTGTCCCCGGCTGCACCCCTTCTC
TCTTCTCTGTACAGTCTCAGTGTGCTTGTGCCCTGGCAAGGCTGAAGGAGGCCACTCCATCTCAC
CTCGGGGGCTGCCCTCAATGTGGAGTGAACCCAGCCAGATCTGAAGGACATTGGGAGAGGGATGCCAGGAA
GCCCTCATCTCAGCAGCCTGGCTGGCATTCCGAAGCTGACTTTCTATAGGCAATTGTACCTTGTGGAGAA
ATGTGTACACCTCCCCAACCCGATTCACTTTCTCTGTAAAAAATAAAAATAAAACAATAAA
AAAA

FIGURE 86

MRLLVAPLLLAWVAGATATVPVVPWHVPCPPQCACQIRPWYTPRSSYREATTVDCNDLFLTA
VPPALPAGTQTLLLQSNSIVRVDQSELGYLANLTEDLSQNSFSDARDCDFHALPQQLSLHL
EENQLTRLEDHSFAGLASLQELYLNHNQLYRIAPRAFSGLSNLLRLHLSNLLRAIDSRWFE
MLPNLEILMIGGNKVDAILDMDNFRPLANRLSVLAGMNLREISDYALEGLQSLESLSFYDNQ
LARVPRRALEQVPGLKFLDLNKNPLQRVGPGDFANMLHLKELGLNNMEELVSIKFALVNLP
ELTKLDITNNPRLSFIHPRAFHLPQMETLMLNNNALSALHQQTVESLPNLQEVLGHGNPIR
CDCVIRWANATGTRVRFIEPQSTLCAEPPDLQRLPVREVFRMTDHCLPLISPRSFPPSLQ
VASGESMVLHCRALAEPEPEIYWVTPAGLRLTPAHAGRRYRVYPEGTLELRRVTAAEAGLYT
CVAQNLVGADETKTVVVGRALLQPGRDEGQGLELRVQETHPYHILLSWTPPNTVSTNLTW
SSASSLRGQGATALARLPRGTHSYNITRLLQATEYWACLQVAFADAHTQLACVWARTKEATS
CHRALGDRPGLIAILALAVLLAAGLAHLGTGQPRKGVGRRPLPPAWAFWGWSAPSVRVV
SAPLVLWPWNPGRKLPRSSEGETLLPPLSQNS

Signal sequence:

amino acids 1-18

Transmembrane domain:

amino acids 629-648

N-glycosylation site.

amino acids 94-98, 381-385, 555-559, 583-587

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 485-489

Casein kinase II phosphorylation site.

amino acids 46-50, 51-55, 96-100, 104-108, 130-134, 142-146,
243-247, 313-317, 488-492, 700-704

Tyrosine kinase phosphorylation site.

amino acids 532-540

N-myristoylation site.

amino acids 15-21, 493-499, 566-572

Amidation site.

amino acids 470-474, 660-664, 692-696

FIGURE 87

GCAAGCCAAGGCCTGTTGAGAAGGTGAAGAAGTCCGGACCCATGTGGAGGAGGGGGACATTGTGTACCGCCT
 CTACATGGCAGACCATCATCAAGGTGATCAAGTTCATCCTCATCATCTGCTACACCGTCTACTACGTGACAA
 CATCAAGTTCGACGTGGACTGCACCGTGGACATTGAGAGCCTGACGGGCTACCGCACCTACCGCTGTGCCACCC
 CCTGGCCACACTCTTCAGATCTGGCTCCTCTACATCAGCCTAGTCATCTTCTACGGCCTCATCTGCATGTA
 CACACTGTGGTGGATGTCACGGCCTCCCTCAAGAAGTACTCGTTGAGTCGATCCGTAGGAGAGCAGCTACAG
 CGACATCCCCGACGTCAAGAACGACTTCGCCTCATGCTGCACCTCATTGACCAATACGACCCGCTCTACTCCAA
 GCGCTTCGGCGTCTCTGTCGGAGGTGAGTGAGAACAGCTGGCAGCTGAACCTCAACAAACGAGTGGACGCT
 GGACAAAGCTCCGGCAGCGCTCACCAAGAACGCGCAGGACAAGCTGGAGCTGCACCTGTTCATGCTCAGTGGCAT
 CCCTGACACTGTGTTGACCTGGAGCTGGAGGTCTCAAGCTGGAGCTGATCCCGACGTGACCATCCCGC
 CAGCATTGCCAGCTCACGGCCTCAAGGAGCTGTGGCTCTACCACACAGCGGCAAGATTGAAGCGCCTGCGCT
 GGCCTTCCTGCGCGAGAACCTGCGGGCGTGCACATCAAGTTCACCGACATCAAGGAGATCCCGCTGTGGATCTA
 TAGCCTGAAGACACACTGGAGGAGCTGCACCTGACGGGCAACCTGAGCGGGAGAACAAACCGCTACATCGTCATCGA
 CGGGCTGGGGAGCTCAAACGCCCTCAAGGTGTCGGGCTCAAGAGCAACCTAACGCAAGCTGCCACAGGTGGTCAC
 AGATGTGGCGTGCACCTGCAGAACGACTGTCCATCAACAAATGAGGGCACCAAGCTCATCGTCTCAACAGCCTCAA
 GAAGATGGCAACCTGACTGAGCTGGAGCTGATCCGCTGGACCTGGAGCGCATCCCCCCTCCATCTTCAGCCT
 CCACAAACCTGAGGAGATTGACCTCAAGGACAACAACCTCAAGAACCATCGAGGAGATCATCGAGCTCCAGCACCT
 GCACCGCTCACCTGCTTAAGCTGTGGTACAACCACATGCCCTACATCCCCATCCAGATCGGCAACCTCACCAA
 CCTGGAGGCCCTCACCTGACCGCAACAAACCTGACCTCCCTCGCGACATCGGCCCTCGCAGAACCTCCAGAACCT
 CTACCTGGACCTCAGCCACAACCGATCGAGACGCTCCCTGGAGCTTCCAGTGCAGGAGCTGAGCTCCAGCACCT
 AGCCATCACGGCCAACCGGATCGAGACGCTCCCTGGAGCTTCCAGTGCAGGAGCTGAGCTCCAGCACCT
 GGGCAACAAACGTGCTGAGTCACTGCCCTCCAGGGTGGCGAGCTGACCAACCTGACGAGATCGAGCTGCCAG
 CAACCGCTGGAGTGCCTGCTGGAGCTGGCGAGTGCCTGAGGAGCTGCTCAAGCGCAGGGCTTGGTGGAGGA
 GGACCTGTTAACACACACTGCCACCCGAGGTGAAGGAGCGGCTGTGGAGGGCTGACAAGGAGCAGGCCTGAGCGAG
 GCCGGCCAGCACAGCAAGCAGCAGGACCGCTGCCAGTCTCAGGCCGGAGGGCTAGCTTCTCCAG
 AACTCCCGACAGCCAGGACAGCTCGGGCTGGCAGGAGCTGGGGCCCTTGAGTCAAGTGCAGGCCAGAGCGAGA
 GGACAGTATCTGTGGGCTGGCCCCCTTTCTCCCTCTGAGACTCACGCCCCAGGGCAAGTGCCTGTGGAGGAG
 AGCAAGTCTCAAGAGCGCAGTATTGGATAATCAGGGCTCCCTGGAGGGCTAGCTCTGCCAGGGCTGAG
 CTGCCACCAAGGGCTGGAGCTGCCCTTCAGTTAGTTCTGGTATTATTTCTCCATCTCCACCTCCTTCATCC
 AGATAACTTATACATTCCAAGAACGTTAGCCCAGATGGAAGGTGTTAGGGAAAGGTGGCTGCCCTTTCCCC
 TTGTCCTTATTAGCGATGCCGCCGGCATTTAACACCCACCTGGACTTCAGCAGAGTGGTCCGGGCAACCAG
 CCATGGGACGGTCACCCAGCAGTGCCGGCTGGCTGCGGTGCCAGGGAGAGCAGGCCCTCAGCTGGAG
 AAGGCCAGGCCCTGGAGCTGCCCTTCAGTTAGTTCTGGCAGTTAGTTAGTTGGTTTTTTTTAACTAAA
 AAAACAATTTTTTAAAAAGCTTGAAGAACGTTAGGGTATTAAAGGAAAGGAAAGGAAAGGAAAGGAAAGGAA
 AAAAGACACTAACGCCAGTGAAGTGGAGTCTCAGGGCAGGGTGGCAGTTCCCTTGAGCAAAGCAGGCCAGACGT
 TGAACTGTTGCTTCCCTGGGAGGGAGGTTTTGGTTGGTTGGTTGGTTGGTCTTGGTGTGACCTTGGTCCAGGAGTT
 CTATTGTTCTGGGAGGGAGGTTTTGGTTGGTTGGTTGGTTGGTTGGTTGGTTGGTCTTGGTGTGACCTTGGTCCAGGAGTT
 ATGTGTCCTGGCAGGCACTCATTTCTGTGGCTGCGCCAGAGGAATGTTCTGGAGCTGCCAAGGAGGGAGGAG
 ACTCGGGTTGGCTAATCCCGGATGAACGGTGTCCATTGCACCTCCCTCGCCTGCCCTGCCCTC
 CGCACAGTCTAACAGGAGCCAGGAGGCCACTCGCCAGACTTTGTTCCCTGCCCTCGCAGTGGGTGT
 CCAGTGCACCGCTGCCCTCCGCTGCTTCCATGCCCTGTCGCCACCTGGTCCCTCATGAAGAGCAGACACTTA
 GAGGCTGGCGGAATGGGGAGGTGCCCTGGAGGGCAGGCAGTGGTGTGGTCTGGAGCCAAACCTGCTTGGAG
 CTGGAGTGCACACAGCCCAGTCGGCACCTGGTGGCTGGAGGCCAAACCTGCTTGGAGCCAAACCTGCTTGGAG
 AGAAGGGTCCCCGCCCTAGATCAATCACGTGGACACTAAGGCACGTTAGAGTCTCTGTCTTAATGATTATGT
 CCATCCGCTGTCCGCTCCATTGTTCTGGCTGTCATTGGATATAATCCTCAGAAATAATGCACACTAG
 CCTCTGACAACCATGAAGCAAAATCCGTTACATGTGGCTGAACCTGTAGACTCGGTACAGTATCAAATAAA
 ATCTATAACAGAAAAAAAAAAAAAA

FIGURE 88

MRQTIIKVIKFILIIICYTVYYVHNIFDVDCVDIESLTGYRTYRCAHPLATLFKILASFYI
SLVIFYGLICMYTLWWMLRRSLKKYSFESIREESSYSDIPDVKNDFAFMLHLIDQYDPLYSK
RFAVFLSEVSENKLRLQNLNNNEWTLKLRQRLTKNAQDKLELHLFMLSGIPDTVFDLVELEV
LKLELIPDVTIPPSIAQLTGLKELWLYHTAAKIEAPALAFLRENLRALHIKFTDIKEIPLWI
YSLKTLEELHTGNLSAENNRYIVIDGLRELKRLKVLRLKSNLSKLPQVVTDVGVHLQKLSI
NNEGTKLIVLNSLKKMANLTELELIRCDLERIPHSIFSLHNLQEIDLKDNNLKIEEIISFQ
HLHRLTCLKLWYNHIAYIPIQIGNLTNLERLYLNRNKIEKIPTQLFYCRKLRYLDLSHNNLT
FLPADIGLLQNLQNLAITANRIETLPPPELFQCRKLRALHLGNNVLQSLPSRVGEINTNLTQIE
LRGNRLECLPVELGECPLLKRSGLVVEEDLFNTLPPEVKERLWRADKEQA

Transmembrane domain:

amino acids 51-75 (type II)

N-glycosylation site.

amino acids 262-266, 290-294, 328-332, 396-400, 432-436, 491-495

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 85-89

Casein kinase II phosphorylation site.

amino acids 91-95, 97-101, 177-181, 253-257, 330-334, 364-368,
398-402, 493-497

N-myristylation site.

amino acids 173-179, 261-267, 395-401, 441-447

FIGURE 89

GCCTGTTGCTGATGCTGCCGTGCGGTACTTGTCATGGAGCTGGCACTGCGCGCTCTCCGT
CCCGCGGTGGTTGCTGCTGCCGTGCTGGCCTGAACGCAGGAGCTGTCATTGACT
GGCCCACAGAGGAGGGCAAGGAAGTATGGATTATGTGACGGTCCGCAAGGATGCCTACATG
TTCTGGTGGCTCTATTATGCCACCAACTCCTGCAAGAACTTCTCAGAACTGCCCTGGTCAT
GTGGCTTCAGGGCGGTCCAGGCAGTTCTAGCACTGGATTGGAAACTTGAGGAAATTGGC
CCCTTGACAGTGATCTAAACACGGAAAACCACCTGGCTCCAGGCTGCCAGTCTCCTATTT
GTGGATAATCCCGTGGCACTGGGTTCAAGTGATGGTAGTGGTGCCTATGCCAAGGA
CCTGGCTATGGTGGCTTCAGACATGATGGTCTCCTGAAGACCTTCTCAGTTGCCACAAAG
AATTCCAGACAGTCCATTCTACATTTCAGAGTCCTATGGAGGAAAATGGCAGCTGGC
ATTGGTCTAGAGCTTATAAGGCCATTCAAGCGAGGGACCATCAAGTGCAACTTGCAGGGGT
TGCCTTGGGTGATTCTGGATCTCCCTGTTGATTGGTCTCTCCTGGGACCTTACCTGT
ACAGCATGTCTCTCGAAGACAAAGGTCTGGCAGAGGTGTCAAGGTTGCAGAGCAAGTA
CTGAATGCCGTAAATAAGGGCTCTACAGAGAGGCCACAGAGCTGTGGGGAAAGCAGAAAT
GATCATTGAACAGAACACAGATGGGTGAACTTCTATAACATCTTAACACTAAAGCACTCCA
CGTCTACAATGGAGTCGAGTCTAGAATTCACACAGGCCACCTAGTTGTCTTGTCAAGC
CACGTGAGACACCTACAACGAGATGCCCTAAGCCAGCTCATGAATGCCCATCAGAAAGAA
GCTAAAATTATTCTGAGGATCAATCCTGGGAGGCCAGGCTACCAACGTCTTGTGAACA
TGGAGGAGGACTTCATGAAGCCAGTCATTAGCATTGTGGACGAGTTGCTGGAGGCAGGGATC
AACGTGACGGTGTATAATGGACAGCTGGATCTCATCGTAGATACCATGGGTCAAGGAGC
GGTGCAGAAACTGAAGTGGCAGAACTGCCTAAATTCAAGTCAGCTGAAGTGGAAAGGCC
ACAGTGCCTAAATCTTGGAAACATCTGCTTGTCAAGTCCTACAAGAACCTTGCTT
TACTGGATTCTGAAAGCTGGTCATATGGTCTCTGACCAAGGGACATGGCTCTGAAGAT
GATGAGACTGGTGACTIONAGGATGGATGGGCTGGAGATGAGCTGGTTGGC
TGGGGCACAGAGCTGAGCTGAGGCCGCTGAAGCTGTAGGAAGCGCCATTCTCCCTGTATCT
AACTGGGCTGTGATCAAGAAGGTTCTGACCAAGCTCTGCAGAGGATAAAATCATTGCTCT
GGAGGCAATTGGAAATTATTCCTGCTTCTTAAAAACCTAAGATTTTAAAAATTGAT
TTGTTTGATCAAAATAAAGGATGATAATAGATATTAA

FIGURE 90

MELALRRSPVPRWLLLLPLLGlnAGAVIDWPTEEGKEVWDYVTVRKDAYMFWWLYYATNSC
KNFSELPLVMWLQGGPGGSSTGFGNFEIGPLSDLKPRKTTWLQAASLLFVDNPVGTGFSY
VNGSGAYAKDLAMVASDMMVLLKTFFSCHKEQTVPFYIFSESYGGKMAAGIGLELYKAIQR
GТИКCNFAGVALGDSWISPVDVLSWGPYLYSMSLLEDKGLAEVSKVAEQVILNAVNGLYRE
ATELGKAEMIIEQNTDGVNFYNILTGSTPTSTMESLEFTQSHLVCLCQRHVRHLQRDALS
QLMNGPIRKKLKIIIPEDQSWGGQATNVFVNMEEDFMKPVISIVDELLEAGINVTVYNGQLDL
IVDTMGQEAWVRKLKWPELPKFSQLKWKALYSDPKSLETSAFVKSYKNLAFYWILKAGHMVP
SDQGDMALKMMRLVTQQE

Signal sequence:

amino acids 1-25

N-glycosylation site.

amino acids 64-68, 126-130, 362-366

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 101-105

Casein kinase II phosphorylation site.

amino acids 204-208, 220-224, 280-284, 284-288, 351-355, 449-453

N-myristoylation site.

amino acids 22-28, 76-82, 79-85, 80-86, 119-125, 169-175,
187-193, 195-201, 331-337, 332-338, 360-366

FIGURE 91

GGCCGCGGGAGAGGAGGCCATGGCGCGCGCGGGCGCTGCTGCTGGCGCTGCTGGCTC
GGGCTGGACTCAGGAAGCCGGAGTCGCAGGAGGCAGCGCCGTTATCAGGACCATGCGGCCGA
CGGGTCATCACGTCGCGATCGTGGGTGGAGAGGACGCCGACTCGGGCGTTGGCGTGGCA
GGGGAGCCTGCGCCTGTGGGATTCCCACGTATCGGAGTGAGCCTGCTCAGCCACCGCTGGG
CACTCACGGCGCGCACTGCTTGAAACCTATAGTGACCTTAGTGATCCCTCCGGGTGGATG
GTCCAGTTGGCCAGCTGACTTCCATGCCATCCTCTGGAGCCTGCAGGCCTACTACACCCG
TTACTTCGTATCGAATATCTATCTGAGCCCTCGCTACCTGGGAATTCACCTATGACATTG
CCTTGGTGAAGCTGTCTGCACCTGTACCTACACTAAACACATCCAGCCCCTGTCTCCAG
GCCTCCACATTGAGTTGAGAACCGGACAGACTGCTGGGTGACTGGCTGGGGTACATCAA
AGAGGATGAGGCACTGCCATCTCCCCACACCCCTCCAGGAAGTTCAAGGTGCCATCATAAACAA
ACTCTATGTGCAACCACCTCTCCTCAAGTACAGTTCCGCAAGGACATCTTGGAGACATG
GTTTGTGCTGGCAACGCCAAGGCGGGAGGATGCCTGCTCGGTGACTCAGGTGGACCCCTT
GGCCTGTAACAAGAATGGACTGTGGTATCAGATTGGAGTCGTGAGCTGGGAGTGGCTGTG
GTCGGCCAATCGGCCGGTGTACACCAATATGCCACCAACTTGGAGTGGATCCAGAAG
CTGATGGCCCAGAGTGGCATGTCCCAGCCAGACCCCTGCCACTACTCTTTCCCTCT
TCTCTGGCTCTCCACTCCTGGGCCGGTTGAGCCTACCTGAGCCATGCAGCCTGGGC
CACTGCCAAGTCAGGCCCTGGTCTCTGTCTTGTGTTGTAATAAACACATTCCAGTTGA
TGCCTTGCAGGGCATTCTTCAAAAAAAAAAAAAAAA

FIGURE 92

MGARGALLLALLLARAGLRKPESQEAAPLSGPCGRRVITSRIVGGEDAELGRWPWQGSLRLW
DSHVCGVSLLSHRWALTAAHCFETYSDSLSDPSGWMVQFGQLTSMPSFWSLQAYYTRYFVSNI
YLSPRYLGNSPYDIALVKLSAPVTYTKHIQPICLQASTFEFENRTDCWVTGWWGYIKEDEALP
SPHTLQEJVQVAIINNSMCNHLFLKYSFRKDIFGDMVCAGNAQGGKDACFGDGGPLACNKNG
LWYQIGVVSGVGCGRPNRPGVYTNISHFEWIQKLMAQSGMSQPDPSWPLLFFPLLWALPL
LGPV

Signal sequence:

amino acids 1-18

N-glycosylation site.

amino acids 167-171, 200-204, 273-277

Casein kinase II phosphorylation site.

amino acids 86-90, 134-138, 161-165, 190-194, 291-295

N-myristoylation site.

amino acids 2-8, 44-50, 101-107, 225-231, 229-235, 239-245,
259-265, 269-275

Amidation site.

amino acids 33-37

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 252-263,

Serine proteases, trypsin family, histidine active site.

amino acids 78-84

FIGURE 93

CCCACGCGTCCGGACGGTGGGAAGGGCAGAATGGGACTCCAAGCCTGCCTCCTAGGGCT
 CTTGCCCTCATCCTCTGGCAAATGCAGTTACAGCCCGAGCCGACCAGCGGAGGACGC
 TGCCCCCAGGCTGGGTGTCCCTGGCCGTGCGAACCTGAGGAAGAGCTGAGTCTCACCTT
 GCCCTGAGACAGCAGAATGTGAAAGACTCTGGAGCTGGTGCAGGCTGTGTCGGATCCCAG
 CTCTCCTCAATACGGAAAATACCTGACCCCTAGAGAAATGTGGCTGATCTGGTGCAGGCCATCCC
 CACTGACCCCTCCACACGGTGCAAAATGGCTCTGGCAGCCGGAGCCAGAAGTGCCTTCT
 GTGATCACACAGGACTTCTGACTTGCTGGTGCAGCATCCGACAAGCAGAGCTGCTGCTCCC
 TGGGGCTGAGTTCATCACTATGTGGGAGGACCTACGGAAACCCATGTTGTAAGGTCCCCAC
 ATCCCTACCAAGCTCCACAGGCCTGGCCCCCATGTGGACTTTGTGGGGACTGCACCGT
 TTTCCCCAACATCATCCCTGAGGCAACGTCTGAGCCGCAGGTGACAGGGACTGTAGGCCT
 GCATCTGGGGTAACCCCTCTGTGATCCGTAAGCGATAACAATTGACCTCACAAGACGTGG
 GCTCTGGCACCAGCAATAACAGCAAGCCTGTGCCCAGTCCCTGGAGCAGTATTCCATGAC
 TCAGACCTGGCTCAGTTCATGCGCCTCTCGGTGGCAACTTGACATCAGGCATCAGTAGC
 CCGTGTGGTTGGACAACAGGGCCGGGCCGGGATTGAGGCCAGTCTAGATGTGCAGT
 ACCTGATGAGTGTGGTCCAACATCTCACCTGGGTCTACAGTAGCCCTGGCCGGCATGAG
 GGACAGGAGCCCTTCCTGCACTGGCTCATGCTGCTCAGTAATGAGTCAGCCCTGCCACATGT
 GCATACTGTGAGCTATGGAGATGATGAGGACTCCCTCAGCAGCCTACATCCAGCGGGTCA
 ACACGTGAGCTCATGAAGGCTGCCGCTCGGGGTCTCACCCGCTCTCGCCTCAGGTGACAGT
 GGGGCCGGTGTGGTCTGTCTCTGGAAAGACACCAGTCCGCCCTACCTTCCCTGCCTCCAG
 CCCCTATGTCAACCACAGTGGAGGCACATCTCCAGGAACCTTCTCATACAAATGAAA
 TTGTTGACTATATCAGTGGTGGTGGCTTCAGCAATGTGTTCCACGGCCTCATACCAAGGAG
 GAAGCTGTAACGAAGTCCCTGAGCTCTAGCCCCCACCTGCCACCATCCAGTTACTTCATGC
 CAGTGGCCGTGCCCTACCCAGATGTGGCTGCACCTTGATGGCTACTGGGTGGTCAGCAACA
 GAGTGCCATTCCATGGGTGTCCGGAACCTCGGCCTCTACTCCAGTGTGTTGGGGATCCTA
 TCCTTGATCAATGAGCACAGGATCCTAGTGGCCCCCCCTCTGGCTTCTCAACCCAAG
 GCTCTACCAGCAGCATGGGCAGGTCTTTGATGTAACCCGTGGCTGCCATGAGTCCGTGTC
 TGGATGAAAGAGGTAGAGGGCCAGGGTTCTGCTCTGGTCTGGATCCTGTAACAGGC
 TGGGGAAACACCAACTTCCAGCTTGCTGAAGACTCTACTCAACCCCTGACCCCTTCCATC
 AGGAGAGATGGCTTGTCCCCCTGCCCTGAAGCTGGCAGTTCACTGCTTATTCTGCCCTGTTG
 GAAGCCCTGCTGAACCCCTCAACTATTGACTGCTGCAGACAGCTTATCTCCCTAACCTGAAA
 TGCTGTGAGCTTGACTCCAAACCTACCATGCTCCATCATACTCAGGTCTCCCTACT
 CCTGCCCTAGATTCTCAATAAGATGCTGTAACTAGCATTTTGATGAAATGCCCTCCCTCCGC
 ATCTCATCTTCCTTTCAATCAGGCTTCAAAAGGGTTGTATACAGACTCTGTGCACTA
 TTCACTTGATATTCAATTCCCAATTCACTGCAAGGAGACCTACTGTCACCGTTACTCT
 TTCCCTACCCCTGACATCCAGAAACAATGGCCTCAGTGCACATCTCAATCTTGTGTTATG
 GCCTTCCATCATAGTTGCCACTCCCTCTCTTACTTAGCTTCCAGGTCTTAACCTCTG
 ACTACTCTTGCTTCCCTCTCATCAATTCTGCTTCTCATGGAATGCTGACCTTCATTGC
 TCCATTGTAGATTTGCTTCTCAGTTACTCATTGCTCCCTGGAACAAATCACTGACA
 TCTACAACCATTACCATCTCACTAAATAAGACTTTCTATCCAATAATGATTGATACTCAA
 TGTA

FIGURE 94

MGLQACLLGLFALILSGKCSYSPEPDQRRTLPPGVSLGRADPEEELSLTFALRQQNVERLS
ELVQAVSDPSSPQYGYLTLENVADLVRPSPLTLHTVQKWLLAAGAQKCHSVITQDFLTCWL
SIRQAELLLPGAEFHYYVGGPTETHVVRSPHPYQLPQALAPHVDFVGGLHRFPPTSSLRQRP
EPQVTGTVGLHLGVTPSVIRKRYNLTSQDVSGTSNNSQACAQFLEQYFHDSDLAQFMRLFG
GNFAHQASVARVVGQQGRGRAGIEASLDVQYLMMSAGANISTWVYSSPGRHEGQEPFLQWLML
LSNESALPHVHTVSYGDDDSLSSAYIQRVNTELMKAAARGLTLLFASGDSGAGCWSVSGRH
QFRPTFPASSPYVTTVGGTSFQEPFLITNEIVDYISGGGFSNVFPRPSYQEEAVTKFLSSSP
HLPPSSYFNASGRAYPDVAALSDGYWVVSNRVPIPWVSGTSASTPVFGGILSLINEHRILSG
RPPLGFLNPRLYQQHGAGLFDVTRGCHESCLDEEVEGQGFCSGPGWDPVTGWGTPTSQLC

Signal sequence:

amino acids 1-16

N-glycosylation site.

amino acids 210-214, 222-226, 286-290, 313-317, 443-447

Glycosaminoglycan attachment site.

amino acids 361-365, 408-412, 538-542

Casein kinase II phosphorylation site.

amino acids 212-216, 324-328, 392-396, 420-424, 525-529

N-myristoylation site.

amino acids 2-8, 107-113, 195-201, 199-205, 217-223, 219-225,
248-254, 270-276, 284-290, 409-415, 410-416, 473-479, 482-488,
521-527, 533-539, 549-555

FIGURE 95

CCCGCGCGCTCTCTCCGGCGCCACACCTGTCTGAGCGGCGCAGCGAGCCGGCCGGC
GGGCTGCTCGGCGCGAACAGTGCTCGGATGGCAGGGATTCCAGGGCTCCTCTTCCTCTC
TTCTTTCTGCTCTGTGCTGTTGGCAAGTGAGCCCTACAGTGCCCCCTGGAAACCCACTTG
GCCTGCATACCGCCTCCCTGCGTCTGCCCCAGTCTACCCCTCAATTAGCCAAGCCAGACT
TTGGAGCCGAAGCCAAATTAGAAGTATCTTCTCATGTGGACCCAGTGTCTATGCCAATGG
CCACTGCCCACTTACGAAGAGGCCAAGCAATATCTGTCTTATGAAACGCTCTATGCCAATGG
CAGCCGCACAGAGACGCAGGTGGCATCTACATCCTCAGCAGTAGTGGAGATGGGCCAAC
ACCGAGACTCAGGGTCTTCAGGAAAGTCTCGAAGGAAGCGGCAGATTATGGCTATGACAGC
AGGTTCAGCATTGGAAAGGACTTCCTGCTCAACTACCCCTTCTAACATCAGTGAAGTT
ATCCACGGGCTGCACCGGACCCCTGGTGGCAGAGAACGATGTCCTCACAGCTGCCACTGCA
TACACGATGGAAAAAACCTATGTGAAAGAACCCAGAAGCTCGAGTGGCTTCTAAAGCCC
AAGTTAAAGATGGTGGTCGAGGGGCCAACGACTCCACTTCAGCCATGCCGAGCAGATGAA
ATTTCACTGGATCCGGGTGAAACGCACCCATGTGCCAAGGGTTGGATCAAGGGCAATGCCA
ATGACATCGGCATGGATTATGATTATGCCCTCTGAACTCAAAAGCCCCACAAGAGAAAA
TTTATGAAGATTGGGTGAGCCCTCTGCTAAGCAGCTGCCAGGGGCCAGGGTCTGGGTCTAT
TGGTTATGACAATGACCGACCAGGCAATTGGTGTATGCTCTGTGACGTCAAAGACGAGA
CCTATGACTTGCTCTACCAGCAATCGATGCCAGCCAGGGGCCAGGGTCTGGGTCTAT
GTGAGGATGTGGAAGAGACAGCAGCAGAAGTGGAGCGAAAAATTATTGGCATTTCAGG
GCACCAGTGGTGGACATGAATGGTCCCCACAGGATTCAACGTGGCTGTCAAATCACTC
CTCTCAAATATGCCAGATTGCTATTGATTAAAGGAAACTACCTGGATTGTAGGGAGGG
TGACACAGTGTCCCTCTGGCAGCAATTAAAGGTCTTCATGTTCTTATTAGGAGAGGCC
AAATTGTTTTGTCTTGCATTGGCGTGACACGTGTGTGTGTGTGTGTGTGTAAAGGTGT
CTTATAATCTTTACCTATTCTACAATTGCAAGATGACTGGCTTACTATTGAAAATG
GTTTGTGTATCATATCATATCATTTAAGCAGTTGAAGGCATACTTGCATAGAAATAA
AAAAAAATACTGATTGGGCAATGAGGAATTGACAATTAAAGTTAATCTCACGTTTG
CAAACTTGATTTCATCTGAACCTGTTCAAAGATTATTAATATTGACATA
CAAGAGATATGAAAAA

FIGURE 96

MAGIPGLLFLFFLLCAVGQVSPYSAPWKPTWPAYRLPVVLPQSTLNLA
KPDFGAEAKLEVS
SSCGPQCHKGTPLPTYEEAKQYLSYETLYANGSRTE
TQVGIYILSSSGDGAQHRD
SGSSGKSR
RRKRQIYGYDSRFSIFGKDFLLNYPF
STSVKLSTGCTGTLVAEKHVL
TAACIHDGKTYVKG
TQKLRVGF
LKPKFKDGGRG
ANDSTS
AMPEQMKF
QWIRVKRTHV
PKGWI
KG
NANDIGMDYD
YAL
LLELKKPHKR
KFMKIGV
SPPAKQL
PGGRIHF
SGYDNDR
PGNLV
YRFCDV
KDETYD
LLYQQCD
AQPGASGSGVY
V
Y
V
RMWKR
QQQKWERK
II
IGIFSGH
QW
V
DMNGSP
QDFNV
AVR
ITPL
KYA
Q
ICYW
IKGNYLDCREG.

Signal sequence:

amino acids 1-19

N-glycosylation site.

amino acids 93-97, 207-211

Glycosaminoglycan attachment site.

amino acids 109-113, 316-320

Casein kinase II phosphorylation site.

amino acids 77-81, 95-99, 108-112, 280-284, 351-355

N-myristoylation site.

amino acids 159-165, 162-168, 202-208, 205-211, 314-320, 338-344

Serine proteases, trypsin family, histidine active site.

amino acids 171-177

FIGURE 97

GCATGCCCTGGGTCTCTGAGCCTGCTGCCTGCTCCCCGCCCCACCAGCCATGGTGGTT
CTGGAGCGCCCCAGCCCTGGTGGGGCTGTCTGGCACCTCACCTCCCTGCTGCTGCTG
GCGTCGACAGCCATCCTCAATGCGGCCAGGATACTGTTCCCCAGCCTGTGGGAAGCCCCA
GCAGCTGAACCGGGTTGTGGCGCGAGGACAGCACTGACAGCGAGTGGCCCTGGATCGTGA
GCATCCAGAAGAATGGGACCCACCCTGCGCAGGTTCTGCTCACCAAGCCGCTGGGTGATC
ACTGCTGCCACTGTTCAAGGACAACCTGAACAAACCATACCTGTTCTGTGCTGCTGGG
GGCCTGGCAGCTGGGAACCCCTGGCTCTGGTCCCAGAAGGTGGGTGTTGCCTGGGTGGAGC
CCCACCCCTGTATTCTGGAAGGAAGGTGCCTGTGCAGACATTGCCCTGGTGCCTCGAG
CGCTCCATACAGTTCTCAGAGCGGGCCTGCCATCTGCCTACCTGATGCCTCTATCCACCT
CCCTCCAAACACCCACTGCTGGATCTCAGGCTGGGAGCAGACATCCAAGATGGAGTTCCCTTGC
CCCACCCCTCAGACCCCTGCAGAAGCTGAAGGTTCTATCATCGACTCGGAAGTCTGCAGCCAT
CTGTACTGGCGGGGAGCAGGACAGGGACCCATCACTGAGGACATGCTGTGCCGGCTACTT
GGAGGGGGAGCGGGATGCTTGTCTGGCGACTCCGGGGCCCCCTATGTGCCAGGTGGACG
GCGCCTGGCTGCTGGCCGGCATCATCAGCTGGGGCGAGGGCTGTGCCGAGCGAACAGGCC
GGGGTCTACATCAGCCTCTGCGCACCGCTGGGTGGAGAAGATCGTGCAGGGGGTGCA
GCTCCCGGGCGCGCTCAGGGGGTGGGCCCTCAGGGCACCGAGCCAGGGCTCTGGGCCG
CCGCGCGCTCCTAGGGCGCAGCGGGACGCGGGCTCGGATCTGAAAGGCGGCCAGATCCACA
TCTGGATCTGGATCTGCGCGGCCTCGGGCGGTTCCCGCCGTAATAGGCTCATCTACC
TCTACCTCTGGGGGCCGGACGGCTGCGGAAAGGAAACCCCTCCCGACCCGCCGAC
GGCCTCAGGCCCCCTCCAAGGCATCAGGCCCCGCCAACGGCCTCATGTCCCCGCCAAC
GACTTCCGGCCCCGGGGCCCCAGCGCTTTGTGTATATAATGTTAATGATTTTAT
AGGTATTGTAACCCCTGCCACATATCTTATTATTCCCTCCAATTCAATAATTATTATT
CTCCAAAAAA

FIGURE 98

```
></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA43318
><subunit 1 of 1, 317 aa, 1 stop
><MW: 33732, pI: 7.90, NX(S/T): 1
MVVSGAPPALGGGCLGTFTSLLLLASTAILNAARI PVPPACGKPQQLNRVVGGEDSTDSEWP
WIVSIQKNGTHHCAGSLLTSRWVITAHCFKDNLNKPYLF SVLLGAWQLGNPGSRQKVGVA
WVEPHPVYSWKEGACADIALVRLERSIQFSEVLPICL PDASIHLPPNTHCWISGWGSIQDG
VPLPHPQTLQKLKVPIIDSEVC SHLYWRGAGQGPITEDMLCAGYLEGERDACLGDGGPLMC
QVDGAWLLAGIISWGEGCAERNRPGVYISLSAHRSWVEKIVQGVQLRGRAQGGGALRAPSQG
SGAAARS
```

Signal sequence:

amino acids 1-32

N-glycosylation site.

amino acids 62-66, 96-100, 214-218, 382-386, 409-413, 455-459,
628-632, 669-673, 845-849, 927-931, 939-943, 956-960

Glycosaminoglycan attachment site.

amino acids 826-830

Casein kinase II phosphorylation site.

amino acids 17-21, 39-43, 120-124, 203-207, 254-258, 264-268,
314-318, 323-327, 347-351, 464-468, 548-552, 632-636, 649-653,
671-675, 739-743, 783-787, 803-807, 847-851, 943-947, 958-962,
1013-1017, 1019-1023, 1021-1025

Tyrosine kinase phosphorylation site.

amino acids 607-615

N-myristoylation site.

amino acids 179-185, 197-203, 320-326, 367-373, 453-459, 528-534,
612-618, 623-629, 714-720, 873-879

FIGURE 99

GACGGCTGGCCACCATGCACGGCTCCTGCAGTTCCCTGATGCTTCGCTGCCGCTACT'GCTA
CTGCTGGTGGCCACCACAGGCCCTGGAGCCCTCACAGATGAGGAGAACGTTGATGGT
GGAGCTGCACAACCTCTACCGGGCCCAGGTATCCCGACGGCCTCAGACATGCTGCACATGA
GATGGGACGAGGAGCTGGCCGCTTCGCAAGGCCTACGCACGGCAGTGCCTGGGGCCAC
AACAAAGGAGCGCGGGCGCCCGGGAGAAATCTGTTGCCATCACAGACGAGGGCATGGACGT
GCCGCTGGCCATGGAGGAGTGGCACCACGAGCGTGAGCACTACAACCTCAGCGCCGCCACCT
GCAGCCCAGGCCAGATGTGCGGCCACTACACGCAGGTGGTATGGGCCAAGACAGAGAGGATC
GGCTGTGGTTCCCACTTCTGTGAGAAGCTCCAGGGTGTGAGGAGACCAACATCGAATTACT
GGTGTGCAACTATGAGCCTCCGGGAACGTGAAGGGAAACGCCCTACCAGGAGGGACTC
CGTGCCTCCAATGTCCTCTGGCTACCACTGCAAGAACTCCCTCTGTGAACCCATCGGAAGC
CCGGAAGATGCTCAGGATTGCCTTACCTGGTAAGTGAGGCCCATCCTTCCGGCGACTGA
AGCATTAGACTCTAGGAAATGGGTACTCCTTCTCCCTAGCAACGGGATTCCGGTTTCT
TGGTAACAGAGGTCTCAGGCTCCCTGGCAACCAAGGCTCTGCCTGCTGTGAAACCCAGGCC
CCAACCTCCTTAGCAACGAAAGACCCGCCCTCATGGCAACAGAGGCTCACCTGCGTAAC
AACTGAGGTCCCTCCATTGGCAGCTCACAGCCTGCCCTTGGATGAGGAGGCCAGTTA
CCTTCCCCAAATGACCCATGTTCTATCCAAAATCAGCAGACAAAGTGACAGACAAAACA
AAAGTGCCCTCTAGGAGGCCAGAGAACTCTCTGGACCCCAAGATGTCCTGACAGGGCAAG
GGAACCTCCTACCCCATGCCAGGAGGAGCTGAGGCTGAGGCTGAGTTGCCTCCTCCAGTG
AGGTCTTGGCCTCAGTTTCCAGGCCAGGACAAGCCAGGTGAGCTGCAGGCCACACTGGAC
CACACGGGCACACCTCCTCCAAGTCCCTGCCAATTCCCAATACCTCTGCCACCGCTAA
TGCCACGGGTGGCGTGCCTGGCTCTGCAGTCGCTTGCAGGTGCAGAGGCCCTGACA
AGCCTAGCGTTGTCAGGGCTGAACCTGGCCCTGGTCAATGTGTGGGCCCTCTCTGGGA
CTACTGCTCCTGCCCTCTGGTGTGGCTGGAATCTCTGAATGGGATACCAACTCAAAGGG
TGAAGAGGTCAAGCTGCTCCTGTCATCTTCCCCACCCCTGCCCCAGCCCTAAACAAGATA
CTTCTTGGTTAAGGCCCTCCGAAGGGAAAGGCTACGGGCATGTGCCTCATCACACCATCC
ATCCTGGAGGCACAAGGCCTGGCTGGCTGCGAGCTCAGGAGGCCCTGAGGACTGCACACC
GGGCCACACCTCCTGCCCTCCCTCTGAGTCCTGGGGTGGGAGGATTGAGGGAGCT
CACTGCCTACCTGGCCTGGGCTGTCCTGCCACACAGCATGTGCCTCTCCCTGAGTGCCTG
TGTAGCTGGGATGGGATTCCTAGGGCAGATGAAGGACAAGCCCCACTGGAGTGGGTTC
TTGAGTGGGGAGGCAGGGACGAGGAAGGAAAGTAACCTGACTCTCCAATAAAACCT
GTCCAACCTGTGAAA

FIGURE 100

MHGSCSFLMLLPLLLLVATTGPVGALTDEEKRLMVELHNLRYRAQVSPASDMLHMRWDEE
LAAFAKAYARQCVWGHNKERGRRGENLFAITDEGMDVPLAMEEWHHEREHYNSAATCSPGQ
MCGHYTQVVWAKTERIGCGSHFCEKLQGVEETNIELVCNYEPPGNVKGRPYQEGTPCSQC
PSGYHCKNSLCEPIGSPEADAQDLPYLVTEAPSFRATEASDSRKMGTPSSLATGIPAFLTEV
SGSLATKALPAVETQAPTSLATKDPPSMATEAPPCVTTEVPSILAHSPLSDEEPVTFPKS
THVPIPKSADKVTDKTKVPSRSPENSLDPKMSLTGARELLPHAQEEAEAEALPPSSEVLAS
VFPAQDKPGELQATLDHTGHTSSKSLPNFPNTSATANATGGRALALQSSLPGAEGPDKPSVV
SGLNSGPGHVWGPLLGLLLLPLVLAGIF

Signal sequence:

amino acids 1-22

N-glycosylation site.

amino acids 114-118, 403-407, 409-413

Glycosaminoglycan attachment site.

amino acids 439-443

Casein kinase II phosphorylation site.

amino acids 29-33, 50-54, 156-160, 195-199, 202-206, 299-303

N-myristoylation site.

amino acids 123-129, 143-149, 152-158, 169-175, 180-186, 231-237,
250-256

Amidation site.

amino acids 82-86, 172-176

Peroxidases proximal heme-ligand signature.

amino acids 287-298

Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7 signature 1.

amino acids 127-138

Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7 signature 2.

amino acids 160-172

FIGURE 101

GTAACTGAAGTCAGGCTTTCACTTGGGAAGCCCCCTCAACAGAATTGGTCATTCTCCAAGTTATGGTGAGCT
ACTCTGTTGTTCTCCCTGCTGCTTTCACATTAGCAGACGGGACTTAAGTCACAACAGATTATCTTTCAT
CAAGGCAGTTCCATGAGCCACCTCAAAAGCCTCGAGAAGTGAACAAACAATGAATTGGAGACCAATTCC
AAATCTGGGACCAGTCTCGGCAAATATTACACTTCTCTTGGCTGGAAACAGGATTGTTGAAATACTCCCTGA
ACATCTGAAAGAGTTCACTGAAACTTGGACCTTAGCAGCAACAATATTCAAGAGCTCAAATCTGATT
TCCAGCCCTACAGCTCAAATATCTGATCTAACAGCAACCGAGTCACATCAATGAACTGGTATTGGACAA
TTTGGCCAACACACTCCTGTTAAAGCTAACAGGAACCGAATCTCAGCTATCCCACCCAAAGATGTTAAACT
GCCCAACTGCAACATCTGAAATTGAAACGAAACAGATTAAAAATGAGATGGACTGACATTCAAGGCTTGG
TGCTCTGAACTCTGAAATGAAAGAAATGGAGTAACGAAACTTATGGATGGAGCTTTGGGGCTGAGCAA
CATGGAAATTTCAGCTGGACCATAACAACTAACAGAGATTACCAAAGGCTGGCTTACGGCTGATGCT
GCAGGAACCTCATCTCAGCCTAACAGGATCAGCCCTGATGCCCTGGAGCTCTGCCAGAACCTCAG
TGAGCTGGACCTAACCTTCAATCACTTATCAAGGTTAGATGATTCAAGCTTCTGGCTAACGTTACTAAATAC
ACTGCACATTGGGAAACAACAGAGTCAGCTACATTGCTGATTGCTGCCCTCCGGGGCTTCCAGTTAAAGACTTT
GGATCTGAAAGAACAAATGAAATTCCCTGGACTATTGAAGACATGAATGGTGTTCTGGGCTTGACAAACTGAG
GCGACTGATACTCCAAGGAAATCGGATCCGTTCTATTACTAAAAAGCCTTCACTGGTTGGATGCATTGGAGCA
TCTAGACTGAGTACAACGCAATCATGCTTACAAGGCAATGCATTTCACAATGAAGAACACTGCAACAAATT
GCATTAAATACATCAAGCCTTTGTCGATTGCCAGCTAAATGGCTCCACAGTGGTGGGGAAAAACAACTT
TCAGAGCTTGTAAATGCCAGTTGCTGCCATCCTCAGCTGCTAAAGGAAGAACGATTGCTGTTAGGCCAGA
TGGCTTGTGTGATGATTTCACAAACCCAGATCACGGTTCAGCAGAACACAGTCGGCAATAAAAGGTTTC
CAATTGAGTTCATGCTCAGCTGCCAGCAGCAGTGATTCCCCAATGACTTTGCTGGAAAAAGAACATGA
ACTACTGCATGATGCTGAAATGGAAATTATGCACACCTCCGGGCCAAGGTGGGAGGTGATGGAGTATACCAC
CATCCTCGGCTGCGCGAGGTGGAATTGCCAGTGAGGGAAATATCAGTGTGTCATCTCAAATCATTGGTT
ATCCTACTCTGTCAAAGCCAAGCTTACAGTAAATATGCTTCCCTATTCAACCAAGACCCCCATGGATCTCACC
CCGAGCTGGGCCATGGCACGCTGGAGTGTGCTGCTGGGGCACCCAGCCCCCAGATAGCCTGGCAGAACGG
TGGGGCACAGACTTCCAGCTGCACGGGAGAGACGCATGCTGATGCCAGGGATGACGTGTTCTTATCGT
GGATGTGAAGATAGAGGACATTGGGGTATACAGCTGCACAGCTCAGAACAGTGCAGGAAGTATTCAAGCAAATGC
AACTCTGACTGTCCTAGAAACACCATCATTTCAGGCCACTGTTGACCGAACTGTAACCAAGGGAGAACAGC
CGTCTACAGTGCATTGCTGGAGGAAGCCCCCTCCCCCTAAACTGAACCTGGACCAAAGATGATAGCCCATTGGTGGT
AACCGAGAGGCACTTTTGCACTGAGGCAATCAGCTCTGATTATTGACTCAGATGTCAGTGATGCTGGAA
ATACACATGTGAGATGCTAACACCCCTGGCACTGAGAGAGGAAACGTGCGCCTCAGTGTGATCCCCACTCAAAC
CTGCGACTCCCCCTCAGATGACAGCCCCATCGTTAGACGATGACGGATGGCCACTGTTGTCGTGATCATAGC
CGTGGTTGCTGTTGGGACGGTCACTCGTGTGGTGTGATCATATACACACAGGGGAGGAATGAAGA
TTGCAGCATTACCAACACAGATGAGACCAACTTGCAGGAGATATTCTAGTTATTGTCATCTCAGGGAACGTT
AGCTGACAGGCAGGATGGGTACGTTGCTTCAAGAAAGTGGAGCCACCCAGTTGTCACATCTCAGGGTGTGG
ATTTTCTTACACACATGACAGTAGTGGACCTGCCATTGACAATAGCAGTGAAGCTGATGTTGGAGCTGC
CACAGATCTGTCCTTGTCCATTGGGATCCACAGGCCATTGATTTGAAGGGAAATGTTGATGGCTCAGA
TCCTTTGAAACATATCATACAGGTTGCACTGCCATTGTCAGGAAAGAACAGTTTAATGGACACTATGAGCCAGTT
CATAAAGAAAAGGAGTGTCTACCTAGATGCCATTGACAATGCGAACGGAGCTCAGTAATATATC
GTGGCCTACATGAGGAAGCTACTAACACTAGTTACTCTCAGAACAGGACCTGGAAATGAAAATCTG
TCTAAACAAGTCTCTTGTGAAATTGACTTGTGCAATCCAGAGGCCAGCGTGTGGCTCGAGTAATTCTT
TACCTTTGGAAAAGCTCTCAGGAGACCTCACCTAGATGCCATTCAAGCTTGGACAGCCATCAGATTGTCAGCC
AAGAGCCTTTATTGAAAGCTCATTCTCCAGACTTGGACTCTGGTCAAGAGGAAGATGGAAAGAACAG
AGATTTTCAGGAAGAAAATCACATTGACCTTAAACAGACTTAACTACAGGACTCAGAACCTTCA
TTATGACTTGGACACATAGACTGAATGAGACAAAGGAAAAGCTTAAACATACTACCTCAAGTGAACTTTT
AAAGAGAGAGAATCTTATGTTTTAAATGGAGTTATGAAATTAAAGGATAAAAATGCTTATTGATACAGAT
GAACCAAATACACAAAGTTATGAAAATTTTAACTGGGAATGATGCTCATATAAGAACACCTTTAAACTA
TTTTTAACCTTGTTTATGCAAAAAGTATCTACGTTAAATGATATAAATCATGATTATTGATTT
TTATAATGCCAGATTCTTTTATGAAAATGAGTTACTAAAGCATTAAATAACCTGCCCTGTGACCA
TTAAATAGAAGTTACTTCATTATTTGACATTATTTAAATGTCAGATTGAA

FIGURE 102

MVDVLLFSLCLLFHISRPDLSHNRLSFIKASSMSHLQLSREVKLNNELETIPNLGPVSAN
 ITLLSLAGNRIVEILPEHLKEFQSLETLDLSSNNISELQTAFPALQLKYLYLNSNRVTSMEP
 GYFDNLANTLLVLKLNRRNISAIPPKMFKLPQLQHLELNRNKKINVDGLTFQGLGALKSLKM
 QRNGVTKLMGAFWGLSNMEILQLDHNNLTEITKGWLYGLLMLQELHLSQNAIRISPAWE
 FCQKLSELDLTFNHLRLSRLDDSSFLGLSLLNTHIGNNRVSYIADCAFRLSSLKTLDDLKNNE
 ISWTIEDMNGAFSGLDKLRRRLILQGNNRIRSITKKAFTGLDALEHLDLSDNAIMSLQGNAFSQ
 MKKLQQLHLNTSSLLCDCQLKWLPOWVAENNQSFVNASCAPQLLKGRSIFAVSPDGVC
 DFPKPQITVQPETQSAIKGSNLNFICSAASSSDSPMTFAWKKDNELLHDAEMENYAHLRAQG
 GEVMEYTTILRLREVEFASEGKYQCVISNHFGSSYSVAKLTVNMLPSFTKTPMDLTIRAGA
 MARLECAAVGHHPAPQIAWQKDGGTDFPAARERRMHVMPEDDVFFIVDVKIEDIGVYSCAQN
 SAGSISANATLTVLETPSFLRPLLDRVTKGAVLQCIAGGSPPPKNWTKDDSPVVTER
 HFFAAGNQLLIVDSDVSDAGKYTCMSNTLGETERGNVRLSVPPTCDSPQMTAPSLLDDG
 WATVGVIIIAVVCVVGTSLVWVVIYHTRRRNEDCSITNTDETNLPADIPSYLSSQGTIAD
 RQDGYVSSESGSHHQFVTSSGAGFFLPQHDSSGTCHIDNSSEADVEAATDLCFLCPFLGSTGP
 MYLKGNVYGSDFETYHTGCSPDPRTVLMHYEPSYIKKECYPCHSEESERSFSNISW
 PSHVRKLLNTSYSHNEPGMKNLCLNKSSLDFSANPEPASVASSNSFMGTFGKALRRPHLDA
 YSSFGQPSDCQPRAFYLKAHSSPDLGSEEDKERTDFQEENHICTFKQTLENYRTPNFQS
 YDLDT

Signal sequence:

amino acids 1-19

Transmembrane domain:

amino acids 746-765

N-glycosylation site.amino acids 62-66, 96-100, 214-220, 382-386, 409-413, 455-459,
 628-632, 669-673, 845-849, 927-931, 939-943, 956-960**Glycosaminoglycan attachment site.**

amino acids 826-830

Casein kinase II phosphorylation site.amino acids 17-21, 39-43, 120-124, 203-207, 254-258, 264-268,
 314-318, 323-327, 347-351, 464-468, 548-552, 632-636, 649-653,
 671-675, 739-743, 783-787, 803-807, 847-851, 943-947, 958-962,
 1013-1017, 1019-1023, 1021-1025**Tyrosine kinase phosphorylation site.**

amino acids 607-615

N-myristoylation site.amino acids 179-185, 197-203, 320-326, 367-373, 453-459, 528-534,
 612-618, 623-629, 714-720, 873-879

FIGURE 103

GGGGAGAGGAATTGACCATGTAAAAGGAGACTTTTTGGTGGTGGCTGTTGGTGCCTTGCAAAATG
 AAGGATGCAGGACGCAGCTTCCTGGAACCGAACGCAATGGATAACTGATTGTGCAAGAGAGAACGGAAAC
 GAAGCTTTCTGTGAGCCCTGGATCTAACACAAATGTATATGTGACACAGGGAGCATTCAAGAATGAAA
 TAAACAGAGTTAGACCCGCGGGGTTGGTGTCTGACATAAATAATCTTAAAGCAGCTGTTCCCTCC
 CCACCCCCAAAAAAAGGATGATTGAAATGAAGAACCGAGGATTACACAAAGAAAAAGTATGTTATTTC
 TATAAAGGAGAAAGTGGAGGAGATTTTGAATGAAAAGTTGGGCTTTTTAGTAAAGTAAAGAAACT
 GGTGTGGTGTTCCTTCTTTGAATTCCCACAAGAGGAGAGGAAATTATAATACATCTGCAAAGAAA
 TTTCAGAGAAGAAAGTTGACCGCGGCAGATTGAGGCATTGATTGGGGAGAGAACCGAGCAGACAGTTGGA
 TTTGTGCTATGTTGACTAAAATTGACGGATAATTGACGGATTGGATTCTTCATCAACCTCCTTTTTAAAT
 TTTTATTCTTTGGTATCAAGATCATGCGTTCTCTGTTCTAACACACCTGGATTCCATCTGGATGTTGCT
 GTGATCAGTCTGAAATACAACCTGTTGAATTCCAGAACGGACCAACACCAAGATAAATTATGAATGTTGAAACAGAT
 GACCTTACATCCACAGCAGATAATGATAGGCTCTAGGTTAACAGGGCCCTATTGACCCCTGCTTGTGGTGCCT
 GCTGGCTCTCAACTTCTGTGGCTGGCTGGTCTGGTGGGGCTCAGACCTGCCCTCTGTGCTCTGAGCAA
 CCAGTTCAAGGAGGATTTGTGTTGGAAAAACCTGGTGGAGGTTCCGGATGGCATCTCCACCAACACAGGCT
 GCTGAACCTCCATGAGAACCAATCCAGATCATCAAAGTGAACAGCAGCTCAAGCAGCTGGAAATCCT
 ACAGTTGAGTAGGAACCATATCAGAACCAATTGAAATTGGGGCTTCAATGGTCTGGCAACCTCAACACTCTGGA
 ACTCTTGACAATGTCCTACTACCATCCGAATTGGAGCTTTGTATACTTGTCTAAACTGAAGGAGCTGGTT
 GCGAAACACCCATTGAAAGCATCCCTTCTATGCTTTAACAGAATTCCCTTTCGCGCCACTAGACTTAGG
 GGAATTGAAAAGACTTCATACATCTCAGAAGGTGCTTGAAGGCTGTCAACTTGAGGTTTTGAACCTTGC
 CATGTGCAACCTCGGGAAATCCCTAACCTCACACCGCTCATAAAACTAGATGAGCTGGATTTCTGGAAATCA
 TTTATCTGCCATCAGGCCTGGCTTTCCAGGGTTTGATGCACCTCAAAACTGTGGATGATACAGTCCAGAT
 TCAAGTGAACGGAATGCCCTTGACAAACCTTCAGTCAGTGGAGATCAACCTGGCACACAATAATCTAAC
 ATTACTGCTCATGACCTCTCACTCCCTGCACTCATCTAGAGCGGATACATTACATCACAAACCTTGGAACTG
 TAACGTGACATACGTGGCTCAGCTGGGATAAAAGACATGGCCCTCGAACACAGCTTGTGCCCCGGTG
 TAACACTCCTCCAACTTAAAGGGAGGTACATTGGAGAGCTGACCAACTTACATGCTATGCTCCGGT
 GATTGTGGAGCCCCCTGCAGACCTCAATGTCAGTGAAGGCATGGCAGCTGAGCTGAAATGTCGGGCTCCACATC
 CCTGACATCTGTATCTGGATTACTCCAAATGGAACAGTCATGACACATGGGGCTACAAAGTGCAGGATAGCTGT
 GCTCAGTGTGGTACGTTAAATTCAACAAATGTAACGTGCAAGATACAGGCTGTACACATGTATGGTGA
 TTCCGTTGGAAATACTACTGCTTCAGCCACCTGAATGTAACGTGCAACCAACTACTCCTTCTTACCTT
 AACCGTCACAGTAGAGACTATGGAACCGTCTCAGGATGAGGCACGGACACAGATAACAATGTGGGCTCCACTCC
 AGTGGTCACTGGGAGACCACCAATGTGACCACTCTCTCACACCACAGAGCACAAGGTGACAGAGAAAACCTT
 CACCATCCCAGTGAATGATATAAACAGTGGGATCCCAGGAATTGATGAGGTATGAAGACTACCAAAATCATCAT
 TGGGTGTTGTGGCATTACACTCATGGCTGAGTGTGCTGGTCAATTCTACAAGATGAGGAAGCAGCACCA
 TCGGCAAAACCATCACGCCAACAAGGACTGTTGAAATTAAATGTGGATGAGATTACGGGAGACACACC
 CATGGAAGCCACCTGCCATGCTCTGCTATGAGCATGAGCACCTAAACTATAACTCATACAAATCTCCCTT
 CAACCACACAACAGTTAACACAATAAATTCAATACACAGTTCACTGATGAACCGTTATTGATCCGAATGAA
 CTCTAAAGACAATGTACAAGAGACTCAAATCTAAACATTACAGAGTTACAAAAACAAACATCAAAAAAA
 GACAGTTTATTAAAATGACACAAATGACTGGGCTAAATCTACTGTTCAAAAAGTGTCTTACAAAAAAACAA
 AAAAGAAAAGAAATTATTATTAAATTCTATTGTGATCTAAAGCAGACAAAAAA

FIGURE 104

MLNKMTLHPQQIMIGPRFNRALFDPLLVLLALQLLVVAGLVRAQTCPSVCSCSNQFSKVIC
VRKNLREVPDGISTNTRLLNLHENQIQIICKVNSFKHLRHEILQLSRNHIRTIEIGAFNGLA
NLNTLELFDNRLTTIPNGAFVYLSKLKELWLRNNPIESIPSYAFNRIPSLRRLDLGELKRLS
YISEGAFEGLSNLRYLNLCNLREIPNLTPLIKDELDSGNHLSAIRPGSFQGLMHLQKL
WMIQSQIQVIERNAFDNLQSLVEINLAHNNLTLPHDLFTPPLHHLERIHLHHNPWCNCIDIL
WLSWWIKDMAPSNTACCARCNTPPNLKGRYIGELDQNYFTCYAPVIVEPPADLNVTGMAAE
LKCRASTSLTSVSWITPNGTVMTHGAYKVRIAVLSGTLNFTNVTVQDTGMYTCMVSNSVGN
TTASATLNVTAATTPFSYFSTVTVETMEPSQDEARTTDNNVGPTVVDWETTNVTTSLTPQ
STRSTEKTFTIPVTDINSGIPGIDEVMKTTKIIIGCFVAITLMAAVMLVIFYKMRKQHHRQN
HHAPTRTVEIINVDEITGDTPMESHLPMPAIEHEHLNHYNSYKSPFNHTTVNTINSIHSS
VHEPLLIRMMNSKDNVQETQI

Signal sequence:

amino acids 1-44

Transmembrane domain:

amino acids 523-543

N-glycosylation site.

amino acids 278-282, 364-368, 390-394, 412-416, 415-419, 434-438,
442-446, 488-492, 606-610

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 183-187

Casein kinase II phosphorylation site.

amino acids 268-272, 417-421, 465-469, 579-583, 620-624

N-myristoylation site.

amino acids 40-46, 73-79, 118-124, 191-197, 228-234, 237-243,
391-397, 422-428, 433-439, 531-537

FIGURE 105

AGCCGACGCTGCTCAAGCTGCAACTCTGTTGAGTTGGCAGTTCTTTCGGTTCCCTCCTGCTGTTGGGGCA
 TGAAAGGGCTTCGCCCGGGAGTAAAGAAGGAATTGACCGGGCAGCGCGAGGGAGGAGCCGCACGCCACCGC
 GAGGGCGGGCGTGCACCCCTGGCTGGAAGTTGTCGCCGGCCCGAGCGCGCAGCCGGCTGGGAGCTTCGGGTAGA
 GACCTAGGCGCTGGACCGCGATGAGCGCCGAGCCTCCGTGCGCGCCGCGGGGTGGGCTGCTGCTGTGC
 CGGGTCTGGGGCGCGCTGGCGGCTCCGACAGCGCGGTGCGGGGAACCTGGGCAGCCCTCTGGGTAGCCGCC
 GAGCGCCCATGCCCACTACCTGCCGCTGCCCTGGGACCTGCTGGACTGCAAGCGGCTAGCGCGTCTT
 CCCGAGCCACTCCCGCTGGTCGCTGGCTGAACTTAAGTCACAACAGATTATCTTCATCAAGGCAAGTTCC
 ATGAGCCACCTTCAAAGCCTTCGAGAAGTGAAGACTGAACAACAATGAATTGGAGACCACTTCAAATCTGGGACCA
 GTCTCGGCAAATATTACACTTCTCTGCTGGAAACAGGATTGTTGAAATACTCCCTGAACATCTGAAAGAG
 TTTCAGTCCCTGAAACTTGGACCTTAGCAGCAACAATATTCAGAGCTCAAACAGTCATTTCCAGCCCTACAG
 CTCAAATATCTGTATCTCAACAGCAACCGAGTCACATCAATGGAACCTGGTATTTGACAATTGGCCAACACA
 CTCCCTGTGTTAAAGCTGAACAGGAACCGAATCTCAGCTATCCCACCCAAGATGTTAAACTGCCAACTGCAA
 CATCTGAATTGAAACCGAAACAAGATTAAAATGTTAGATGGACTGACATTCCAAGGCCCTGGTCTGAAAGTCT
 CTGAAAATGCAAAGAAATGGAGTAACGAAACTTATGGATGGAGCTTTTGGGGCTGAGCAACATGGAATTG
 CAGCTGGACCATAACACCTAACAGAGATTACCAAGGGTGGCTTACGGCTGCTGATGCTGAGGAACCTCAT
 CTCAGCCAAATGCCATCAACAGGATCAGCCTGATGCTGGAGTTCTGCCAGAAGCTCAGTGAGCTGGACCTA
 ACTTTCAATCACTTATCAAGGTTAGATGATTCAAGCTTCTTGGCTAAGCTTACTAAATACACTGCACATTGGG
 AACAAACAGAGTCAGCTACATTGCTGATTGCGCTTCCGGGGCTTCCAGTTAAAGACTTGGATCTGAAAGAAC
 AATGAAAATTCTGGACTATTGAAGACATGAATGGTCTTCTGGCTGACAAACTGAGGGACTGATACTC
 CAAGGAAATCGGATCCGTTCTATTACTAAAAAGCTTCACTGGTTGGATGCAATTGGAGCATCTAGACCTGAGT
 GACAACGCAATCATGCTTCAAGGCAATGCAATTTCACAAATGAAGAAACTGCAACAATTGCAATTAAATACA
 TCAAGCCTTTGTCGATTGCCAGCTAAATGGCTCCACAGTGGGGGGAAACAAACTTTCAGAGCTTGTGA
 AATGCCAGTTGTGCCCATCCTCAGCTGCTAAAGGAAGAAGCATTGCTGTTAGCCCAGATGGCTTGTGT
 GATGATTTCACCAACCCCCAGATCACGGTTAGCCAGAACACAGTCGGCAATAAAAGGTTCAATTGAGTT
 ATCTGCTCAGCTGCCAGCAGCAGTGAATTCCCAATGACTTTGCTTGGAAAAAAAGACAATGAACTACTGCA
 GCTGAAATGAAAATTATGCAACACCTCCGGCCCAAGGTGGCGAGGTGATGGAGTATACCACCATCTCGGCTG
 CGCGAGGTGGAATTGCCAGTGAAGGGAAATATCAGTGTGTCATCTCAACTTTGGTCTCATCCTACTCTGTC
 AAAGCCAAGCTTACAGTAAATATGCTTCCCTCATTCAACAGACCCCCATGGATCTCACCACCGAGCTGGGCC
 ATGGCACCGCTGGAGTGTGCTGTTGGGCCACCCAGCCCCAGATAGCTGGCAGAAGGATGGGGCACAGAC
 TTCCAGCTGCACGGGAGAGACGCATGCTGATGCCAGGATGACGTGTTCTTATCGTGGATGTGAAGATA
 GAGGACATTGGGTATAACAGCTGCACAGCTCAGAACAGTGCAGGAAGTATTCAGCAAATGCAACTCTGACTGTC
 CTAGAACACCATCATTTGCCACTGTTGGACCGAACTGTAACCAAGGGAGAACAGCCGCTTACAGTGC
 ATTGCTGGAGGAAGCCCTCCCTAAACTGAACTGGACAAAGATGATGCCATTGGTGTAAAGGAGACGGC
 TTTTGCAGGCAATCAGCTCTGATTATTGTAAGTCTCAGATGTCAGTGATGCTGGAAATACACATGTGAG
 ATGCTAACACCTTGGCACTGAGAGAGGAAACAGTGCAGGCTCAGTGATCCCCACTCCAACCTGCA
 CAGATGACAGCCCCATCGTAGACGATGGCAGTGGCACTGTGGTGTGATCATAGCCGTGGTTGCTGT
 GTGGTGGGACGTCACTCGTGTGGTGTGATCATATACCACACAAGGCGGAGGAATGAAGATTGCA
 AACACAGATGAGACCAACTGCCAGCAGATATTCTAGTTATTGTCATCTCAGGAAACGTTAGCTGACAGGCAG
 GATGGGTACGTGCTTCAGAAAGTGGAGGCCACCCAGGTTGTCACATCTTCAGGTGCTGGATT
 CAACATGACAGTAGTGGACCTGCCATTGACAATAGCAGTGAAGCTGATGTTGAAAGCTGCCACAGATGTT
 CTTTGTCCGTTTGGGATCCACAGGCCCTATGTTGAAGGGAAATGTGATGGCTCAGATCTTGTAA
 TATCATACAGGTTGCACTGCCAGAACAGTTAATGGACCAACTATGAGCCAGTTACATAAAGAAAAAG
 GAGTGCTACCCATGTTCTCATCTTCAGAACGAACTCTGCAACGGAGCTTCAGTAATATATCGTGGCTTACAT
 GTGAGGAAGCTACTTAACACTAGTTACTCTCACAAATGAAGGACCTGGAATGAAAATCTGTGCTAAACAAGTCC
 TCTTGTGAAATGCAATCCAGAGCCAGCGTCGGTGCCTCGAGTAATTCTTCATGGTACCTTGGAAA
 GCTCTCAGGAGACCTCACCTAGATGCCATTCAAGCTTGGACAGCCATCAGATTGTCAGCAAGAGCCTT
 TTGAAAGCTCATTCTTCCCAGACTGGACTCTGGGTGAGAGGAAGATGGGAAAGAAAGGACAGATTTT
 GAGGAAAGATCATTTGACATTAAACAGACTTAGAAAACACTACAGGACTCCAAATTTCAGTCTTATGACTTGGAC
 ACATAGACTGAATGAGACCAAGGAAAGCTTAAACATACTACCTCAAGTGAACCTTTATTAAAAGAGAGAGA
 ACTTATGTTTTAAATGGAGTTATGAATTAAAGGATAAAAGCTTATTATACAGATGAACCAAAATTAC
 AAAAGTTATGAAAATTGTTACTGGGAATGATGCTCATATAAGAATACCTTTAAACTATTTTA
 ACTTTGTTTATGCAAAAGTATCTTACGTTAAATTATGATATAATCATGATTATTGTTATGTT
 TTTCTTTTATGGAAAATGAGTTACTAAAGCATTAAATAACCTGCCCTGTACCAATT
 TTTAAATAGAAGTTACTTCATTGACATTATTTAATAAAATGTGCAATTGAAAA
 ACTTCATTATATTGACATTATTTAATAAAATGTGCAATTGAAAAA

FIGURE 106

MSAPSLRARAAGLGLLLCAVLGRAGRSDSGGRGELGQPSGVAERPCPTTCRCLGDLDCSR
 KRLARLPEPLPSWVARLDLSHNRLSF IKASSMSHLQSLREVKLNNNELETIPNLGPVSANIT
 LLSLAGNRIVEILPEHLKEFQSLETLDLSSNNISELQTAFPALQLKYLYLNSNRVTSMEPGY
 FDNLANTLLVLKLNRRNRI SAIPPKMFKLPQLQHLELNRRN KIKNVDGLTFQGLGALKSLKMQR
 NGVTKLMGDGFWGLSNMEILQLDHNNLTEITKGWLYGLLMLQELHLSQNAINRISPDAWEFC
 QKLSELDLTFNHLSRLDDSSFLGLSLLNLT HIGNNRVSYIADCAFRLSSLKTLDDLKNNEIS
 WTIEDMNGAFSGLDKLRRRLILQGNRIRSI TKAFTGLDALEHLDLSDNAIMSLQGNAFSQMK
 KLQQLHLNTSSLLCDCQLKWLPOWVAENNQSFVNASC AHPQLLKGRSIFAVSPDGVCDDF
 PKPQITVQPETQSAIKGSNLSFICSAASSSDSPMTFAWKKDNELLHDAEMENYAHLRAQGGE
 VMEYTTILRLREVEFASEGKYQCVISNHFGSSYSVKAKLTVNMLPSFTKTPMDLTIRAGAMA
 RLECAAVGHPAPQIAWQKDGTDFFAARERRMHVMPEDDVFFIVDVKIEDIGVYSCTAQNSA
 GSISANATLTVLETPSFLRPLLDRVTKGETAVLQCIAGGSPPPKNWTKDDSPVVTERHF
 FAAGNQLLIIVDSDVSDAGKYTCEMSNTLGTERGNVRLS VIPTPTCDSPQMTAPS LDDGWA
 TVGVVIIAVVCCVVGTSLVWVVIYHTRRNEDCSITNTDETNLPADIPSYLSSQGT LADRQ
 DGYVSSESGSHHQFVTSSGAGFFLPQHDSSGTCHIDNSSEADVEATDLFLCPFLGSTGPMY
 LKGNVYGSDFETYHTGCPDPRTVLMHYEPSYIKKKECYPCHPSEES CERSFSNISWPS
 HVRKLLNTSYSHNEGPGMKNLCLNKSSLDFSANPEPASVASSNSFMGTFGKALRRPHLDAYS
 SFGQPSDCQPRAFYLKAHSSPDLD\$GSEEDGKERTDFQEENHICTFKQTLENYRTPNFQSYDLDT

Signal sequence:

amino acids 1-27

Transmembrane domain:

amino acids 808-828

N-glycosylation site.

amino acids 122-126, 156-160, 274-278, 442-446, 469-473, 515-519, 688-692, 729-733, 905-909, 987-991, 999-1003, 1016-1020

Glycosaminoglycan attachment site.

amino acids 886-890

Casein kinase II phosphorylation site.

amino acids 99-103, 180-184, 263-267, 314-318, 324-328, 374-378, 383-387, 407-411, 524-528, 608-612, 692-696, 709-713, 731-735, 799-803, 843-847, 863-867, 907-911, 1003-1007, 1018-1022, 1073-1077, 1079-1083, 1081-1085

Tyrosine kinase phosphorylation site.

amino acids 667-675

N-myristoylation site.

amino acids 14-20, 36-42, 239-245, 257-263, 380-386, 427-433, 513-519, 588-594, 672-678, 683-687, 774-780, 933-939

Leucine zipper pattern.

amino acids 58-80, 65-87

FIGURE 107

CAAAACTTGCGTCGCGGAGAGGCCAGCTGACTTGAATGGAAGGAGCCCGAGCCCGCGGAGCGCAGCTGAGAC
 TGGGGGAGCGCGTTCGGCTGTGGGCGCCCTCGGCCGGGGCGCAGCAGGGAAAGGGGAAGCTGTGGCTGCC
 CTGCTCCACAGGGCCCACTGGTGTGAACCGGGAGAGCCCTGGGTGGCTCCCTATCCCTCTTATATA
 GAAACCTTCCACACTGGGAAGGCAGCGCGAGGCAGGGCTCATGGTGAGCAAGGAGGCCGCTGATCTGCAG
 CGCACAGCATTCCGAGTTACAGATTTACAGATACAAATGGAAGGCGAGGAGGCAGAACAGCCTGCC
 TCCATCAGCCCTGGCGCCAGGCATCTGACTCGGCACCCCTGCAGGCACCATGGCCAGAGCCGGTGTGC
 TGCTCCTGCTGCTGCTGCCACAGCTGCACCTGGGACCTGTGCTTGCCTGAGGGCCCCAGGATTTGGCCGA
 GTGGCGGCCACAGCTGAGCCCCGAAGAGAACGAATTGCGGAGGAGGCCGGTGTGGTACTGAGCCCTGAGG
 AGCCCCGGCTGGCCAGCCGCGGTCAAGCTGCCCGAGACTGTGCTGTCCAGGAGGGCGTGTGGACTGTG
 CGGGTATTGACCTGCGTGAGTCCCAGGGACCTGCCTGAGCACACCAACCACCTATCTCTGCAGAACACCAGC
 TGGAAAAGATCTACCCCTGAGGAGCTCTCCCCGCTGCACCGGCTGGAGACACTGAACCTGCAAAACAACC
 CTTCCGAGGGCTCCACAGAGAAGGCAGTTGAGCATCTGACCAACCTCAATTACCTGTACTTGGCCAATAACAAGC
 TGACCTTGGCACCCGCTTCTGCCAACGCCCTGATCAGTGTGGACTTGTGCTGCCAATATCTACCAAGATCT
 ATGGGCTCACCTTGGCCAGAACCAAACCTGGAGTCTGTGACCTGCACAAACAAGCTGGCAGACGCCGGC
 TGCGGACAACATGTCACGGCTCCAGCAACGTCGAGGTCTCATCTGTCCAGCAACCTCTGCACCGTGC
 CCAAGCACCTGCCGCTGCCCTGACAAGCTGCACCTCAAGAACAAAGCTGGAGAAGATCCCCCGGGGCG
 TCAGCGAGCTGAGCAGCCTGCGCAGCTATACTGCAGAACAAACTACCTGACTGACGAGGGCGTGGACAACGAGA
 CCTTCTGGAAGCTCCAGCCTGGAGTACCTGGATCTGTCCAGCAACAACTGTCTGGGTCCAGCTGGCTGC
 CGCGCAGCCTGGTGTGCTGCACCTGGAGAAGAACGCCATCCGGAGCGTGGACCGAATGTGCTGACCCCCATCC
 GCAGCCTGGAGTACCTGCTGCTGCACAGCAACAGCTGCCAGGGCATCCACCCACTGCCCTCCAGGGCC
 TCAAGCGTTGACACGGTGACCTGTACAACAACGCGCTGGAGCGCTGCCAGTGGCTGCCATGTGCTGCGTGC
 GCACCCCTATGATCTGCACAACCAAGATCACAGGATTGGCGCGAAGACTTTGCCACCACTACTCTGGAGG
 AGCTCAACCTCAGCTACAACCGCATCACCAGCCCACAGGTGACCGCAGCCTCCGCAAGCTGCCCTGCTGC
 GCTCGTGGACCTGTCGGCAACGGCTGACACGCTGCCACCTGGCTGCCATGGCTGCTGAAATGTCAATGTGCTGAAGG
 TCAAGCGCAATGAGCTGGCTGCCCTGGCACAGGGCGCTGGCGGGATGGCTCAGCTGCGTGAGCTGTACCTCA
 CCAGCAACCGACTGCCAGCCAGCCCCTGGGCCCCCTGCTGGGACCTGCCCATCTGAGCTGCTGGACA
 TCGCGGGAAATCAGCTACAGAGATCCCCGAGGGCTCCCCGAGTCACCTGAGTACCTGTACTGCCAGAACAAACA
 AGATTAGTGCCTGGCCTGGCACAGTGCCTTCCGGAGGCTGAAGCACCTGAGGTCTGGACATTGAAGGCAACT
 TGGCTGGGCTCCGTGGACAGTGCCTTCCGGAGGCTGAAGCACCTGAGGTCTGGACATTGAAGGCAACT
 TAGAGTTGGTGACATTCAAGGACCGTGGCGCTTGGGAAGGAAAAGGAGGAGGAGGAAGGAGGAGGAGG
 AGGAAGAGGAAACAAGATAGTGAACAGGTGATGCAAGATGTGACCTAGGATGATGGACCGCCGACTCTTCTGC
 AGCACACGCCCTGTCGTGAGCCCCCCTCTGCCATCACAGACACACCCAGCTGCACACATGAGGCA
 TCCCACATGACACGGCTGACACAGTCTCATATCCCCACCCCTTCCACGGCGTGTCCACGGCCAGACACATGC
 ACACACATCACCCCTCAAACACCCAGCTGCCACACACAACTACCCCTCAAACCCACACAGTCTGTACAC
 CCCCACATGCCACGGCTGCCACGCCCTGTAATCATGCAAGGGAGGGCTGCCACACAGGCCACCC
 TTCCCTCCCCCTGCTGACATGTGTATGCGTATGCATACACACCACACACACATGCAAGTCATGTGCGAA
 CAGCCCTCAAAGCTATGCCACAGACAGCTTGGCCAGCCAGAATGCCATAGCAGCTGCCGTGCC
 GTCCATCTGTCGTCGTTCCCTGGAGAAGACACAAGGTATCCATGCTCTGTGCCAGGTGCCACCC
 GGAACTCACAAAGCTGGCTTTATTCTCTTCCATCTATGGGACAGGAGCCTTCAGGACTGCTGGCTGCC
 TGGCCACCTGCTCTCCAGGTGCTGGCAGTCACCTGCTAAGAGTCCCTCCCTGCCACGCCCTGGCAGGACA
 CAGGCACCTTCCAATGGCAAGCCCAGTGGAGGCAGGATGGAGAGGCCCTGGGTGCTGGGCCCTGGGG
 CAGGAGTGAAGCAGAGGTGATGGGCTGGCAGGCCAGGGAGGAAGGACCCAGCTGCACCTAGGAGACACCTT
 GTTCTTCAGGCCTGTTGGGAAGTTCCGGGTGCTTTATTCTTTATTCTTTCTAAGGAAAAAAATGATAAAAT
 CTCAAAGCTGATTTCTTGTATAGAAAACATAATAAAAGATTATCCCTATCCCTGCCAAAAAA

FIGURE 108

MEGEEAEQPAWFHQWPWRPGASDSAPPAGTMAQSRVLLLLLPPQLHLGPVLAVRAPGFGRS
GGHSLSPPEENEFAAEEPVLVLSPEEPGPAAVSCPRDCACSQEGVVDCGGIDLREFPGDLP
EHTNHLSLQNNQLEKIYPEELSRLHRLETNLQNNRLTSRGLPEKAFEHTNLNYLYLANNK
LTLaPRFLPNALISVDFAANYLTKIYGLTFCQKPNLRSVYLHNNKLA DAGLPDNMFNGSSNV
EVLILSSNFLRHVPKHLPPALYKLHLKNNKLEKIPPGAFSELSSLRELYLQNNYLTDEGLDN
ETFWKLSSLEYLDLSSNNLSRVPAGLPRSLVLLHLEKNAIRSVDANVLTPIRSLEYLLLHSN
QLREQGIHPLAFQGLKRLHTVHLYNNALERVPSGLP RVR TLMILHNQITGIGREDFATTYF
LEELNLSYNRITSPQVHRDAFRKLRLRSLDLSGNRLHTLPPGLPRNVHVLKVKRNELAALA
RGALAGMAQLRELYLTSNRLRSRALGPRAWVDLAHLQLLDIAGNQLTIEPEGLPESLEYLYL
QNNKISAVPANAFDSTPNLKGIFLRFNKLAVGSVVDSA FRRLKHLQVLDIEGNLEFGDISKD
RGRLGKEKEEEEEEEEEEETR

Signal sequence:

amino acids 1-48

N-glycosylation site.

amino acids 243-247, 310-314, 328-332, 439-443

Casein kinase II phosphorylation site.

amino acids 68-72, 84-88, 246-250, 292-296, 317-321, 591-595

N-myristoylation site.

amino acids 19-25, 107-113, 213-219, 217-223, 236-242, 335-341,
477-483, 498-502, 539-545, 548-554

Leucine zipper pattern.

amino acids 116-138, 251-273, 258-280, 322-344, 464-486, 471-493,
535-557

FIGURE 109

GGGAGGGGGCTCCGGGCGCCGCGCAGCAGACCTGCTCCGGCCGCGCCTGCCGCTGTCCTCCGGAGCGGCAG
 CAGTAGCCGGGCGGGAGGGCTGGGGTCTCGAGACTCTAGAGGGCGCTCCCACACCC
 CAACCTGTTCTCGCGCCACTGCGCTGCCAGGACCCGCTGCCAACATGGATTTCCTGGCGCTGGT
 GCTGGTATCCTCGCTCACCTGCAAGGGCCGGAGTTGACAGGGAGGGCTGGCCAGGCAAATAGTGTATCGAT
 TGGCCTATGTCGTTATGGTGGGAGGATTGACTGCTGCTGGGCTGGCTGCCAGTCTGGGACAGTGTCA
 TGGTGTGCAACCACGATGCAAACATGGTGAATGTATCGGGCAAACAAGTGAAGTGTATCTGGTTATGCTGG
 AAAAACCTGTAATCAAGATCTAAATGAGTGTGGCTGAAGCCCGCCCTGTAAGCAGGGTGCATGAACACTTA
 CGGCAGCTACAAGTGTACTGTCACGGATATATGCTCATGCCGATGGTCTGCTCAAGTGCCTGACCTG
 CTCCATGGCAAACGTCAAGTATGGCTGTGATGGTTAAGGACAAACGGTGCAGTGCCATCCCCCTGGCCT
 GCACCTGGCTCTGATGGGAGGACTGTGAGATGATGAATGTCTACAGGAAGAGCCTCTGCCCTAGATT
 TAGGCAATGTCACACTTTGGGAGCTACATCTGCAAGTGTCTAAAGGCTTCGATCTCATGTATATTGGAGG
 CAAATATCAATGTCATGACATAGACGAATGTCACTTGGTCAGTATCAGTGCAGCAGCTTGCTCGATGTATAA
 CGTACGGGCTACAAGTGAATGTAAGAAGGATACCGGGTGTGGACTGACTTGTGTATATCCAAA
 AGTTATGATTGAAACCTTCAGGTTCAATTGTAACCAAGGGAAATGGTACCATTTAAAGGGTACACAGGAAA
 TAATAATTGGATTCTGATGTTGGAAGTACTTGGTGGCTCGAAGACACCATAATTCTCTATCATTAACCAA
 CAGGCCTACTTCTAACGCAACAAACAAGACCTACACCAAAAGCCAACACCAATTCTACTCCACCAACCA
 CTCGCCAACAGAGCTCAGAACACCTCTACACCTACAACCCAGAAGGCCAACACGGACTGACAACATATAGC
 ACCAGCTGCCAGTACACCTCCAGGGGATTACAGTTGACAACAGGGTACAGACAGACCCCTCAGAAACCCAGGG
 AGATGTTGTCAGTGTCTGGTACACAGTTGTAATTGGACCATGGACTTGTGGATGGATCAGGGAGAAAGACAA
 TGACTTGCACGGGACCAATCAGGACCCAGCAGGTGACAATATCTGACAGTGTGGCAGGCCAAGGCCAGG
 GGGAAAGCTGCACCTGGTGTACCTCTGGCCCTCATGCAATTGGGACCTGTGCCGTGCAATTAGGCA
 CAAGGTGACGGGCTGCACTGGCACACTCCAGGTGTTGTGAGAAAACAGGTGCCACGGAGCAGCCCTGTG
 GGGAAAGAAATGGTGGCCATGGTGGAGGAAACACAGATCACCTGGAGGGCTGACATCAAGAGCGAATCACA
 AAGATGATTAAAGGGTGGAAAAAAAGATCTATGATGGAAAATTAAAGGAACGGGATTATTGAGCCTGGAGAAG
 AGAAGACTGAGGGCAACCATGATGGTTTCAAGTATATGAGGGTGGCACAGAGAGGGTGGCAGCAGCTG
 TTCTCCATATGCACTAAGAACAGAGGAAACTGGCTTAGACTAGAGTATAAGGGAGCATTCTGGCAGG
 GCCATTGTTAGAATACTTCATAAAAAAAAGATGTTAAAGATGTTCTTACCAAGGAAAGTAACAAATTATGAA
 TAAAAATTGTCATTTAAGATGGTTAAAGATGTTCTTACCAAGGAAAGTAACAAATTATGAAATTCCCAA
 AGATGTTGATCCTACTAGTAGTATGCACTGAGTAAATCTTAAAGGAAACTTTAGAACTAAATAATTG
 GACAAGGCTTAATTAGG
 CATTTCCTCTTGACCCCTAATGGAGAGGGATTGAAAGGGGAGAGCCCACCAATGCTGAGCTCA
 TCTCTCCCTATGGCAATCCTAGCAGTATTAAGAAAAGGAAACTATTATCCAAATGAGAGTATGATGGAC
 AGATATTAGTATCTGTAATGTCCTAGTGTGGCGGTGTTCAATGTTCTCATGGTAAAGGTATAAGGC
 TTTCATTGTTCAATGGATGATGTTCAAGTTTTTTTTAAGAGATCCTCAAGGAACACAGTTAGAGAG
 ATTTTCATCGGGTGCATTCTCTGCTTCGTGTGACAAGTTATCTGGCTGCTGAGAAAAGAGTGCCTGCC
 ACACGGCAGACCTTCCTCACCTCATCAGTATGATTCTCTATCAATTGGACTCTCCAGGTTCCAC
 AGAACAGTAATATTTTTGAACAATAGGTACAATAGAAGGTCTTGTGTTAACCTGGTAAAGGAGGGCTGG
 AGGGGAAATAATCATTAAGGCTTGGAGTAACGGCAGAATATAGGCTGAGATCCATTAAATGGTCATT
 TCCTTATGGTCATATAACTGCAAGCTGAAGATGAAAGGGAAAATAATGAAATTTCCTTCA
 TGATACATTGCACTAAACTGATGGAAGAAGTTATCAAAGTACTGTATAACATCTGTTTATTATTA
 ATGTTCTTAAATGGTAAATAAAACACTGTTAGTAAT
 CTAATAAAATGGTGTGGTTTCAAATGGCTAATAAAACAATTATGTA
 AAATAAAACACTGTTAGTAAT

FIGURE 110

MDFLLALVLVSSLYLQAAAEDGRWPRQIVSSIGLCRYGGRIDCCWGWARQSWGQCQPVVCQP
RCKHGEVICGPNCCKCHPGYAGKTCNQDLNECGLKPRPCKHRCMNTYGSYKCYCLNGYMLMPD
GSCSSALTCSMANCQYGCDVVKGQIRCQCPSPGLHLAPDGRTCDVDECATGRASCPRFRQC
VNTFGSYICKCHKGFDLMLYIGGKYQCHDIDECSLGQYQCSSFARCYNVRGSYKCKCCKEGYQG
DGLTCVYIPKVMIEPSGPIHVPKGNGTILKGDTGNNNWIPDVGSTWWPPKTPYIPPIITNRP
TSKPTTRPTPKPTPIPTPPPPPLPTELRTPLPPTPERPTTGLTTIAPAASTPPGGITVDN
RVQTDPKQKPRGDVFSVLVHSCNFDHGLCGWIREKDNDLHWEPIRDPAQQYLTVSAAKAPGG
KAARLVLPLGRLMHSGDLCLSFRHKVTGLHSGTLQVFVRKHGAHGAALWGRNGGHGWRQTQI
TLRGADIKSEQR

Signal sequence:

amino acids 1-17

N-glycosylation site.

amino acids 273-277

Casein kinase II phosphorylation site.

amino acids 166-170, 345-349

Tyrosine kinase phosphorylation site.

amino acids 199-206

N-myristoylation site.

amino acids 109-115, 125-131, 147-153, 191-197, 221-227, 236-242,
421-427, 433-439, 462-468, 476-482

Aspartic acid and asparagine hydroxylation site.

amino acids 104-116, 186-198, 231-243

Cell attachment sequence.

amino acids 382-385

EGF-like domain cysteine pattern signature.

amino acids 75-87

FIGURE 111

CTTCTTGAAAAGGATTATCACCTGATCAGGTTCTCTGCATTTGCCCTTAGATTGTGA
AATGTGGCTCAAGGTCTTCACAACTTCCCTTGCACAGGTGCTGCTGGGGCTGA
AGGTGACAGTGCCATCACACACTGTCATGGCGTCAGAGGTCAAGGCCCTACCTACCCGTC
CACTATGGCTTCCACACTCCAGCATCAGACATCCAGATCATGGCTATTGAGAGACCCCA
CACAATGCCAAATACTTACTGGCTCTGTGAATAAGTCTGTGGCTTGACTTGAATACC
AACACAAGTTACCATGATGCCACCCAAATGCATCTGCTTATCAACCCACTGCAGTCCCT
GATGAAGGCAATTACATCGTAAGGTCAACATTCAAGGGAAATGGAACACTATCTGCCAGTCA
GAAGATAACAAGTCACGGTTGATGATCCTGTCAAAAGCAGTGGTGAGATTCACTCCTCC
CTGGGGCTGTGGAGTATGTGGGAACATGACCTGACATGCCATGTGAAGGGGGCACTCGG
CTAGCTTACCAATGGCTAAAAAATGGAGACCTGTCCACACCAGCTCCACCTACTCCTTTC
TCCCCAAAACAATACCCTCATATTGCTCCAGTAACCAAGGAAGACATTGGAATTACAGCT
GCCTGGTGGAGAACCTGTCAGTGAAATGGAAGTGAATTGATAAAAGGGCTAAAGTAGGGGAAGTGT
TGACCTTGGAGAGGCCATCCTATTGATTGTTCTGCTGATTCTCATCCCCCAACACCTACT
CCTGGATTAGGAGGAACGACAATACTACATATCATTAAGCATGGCCTCGCTTAGAAGTT
GCATCTGAGAAAGTAGCCCAGAACAGAACATGGACTATGTGTGCTGTACAAACAACATAAC
CGGCAGGCAAGATGAAACTCATTCACAGTTACATCACTCCGTAGGACTGGAGAAGCTG
CACAGAAAAGGAAATCATTGTCACCTTAGCAAGTATACTGGAATATCACTATTTGATT
ATATCCATGTGTTCTCTCCATGGAAAAAATATCAACCTACAAAGTTAAACAGAA
ACTAGAAGGCAGGCCAGAACAGAACATACAGGAAAGCTCAAACATTTCAGGCCATGAAGATG
CTCTGGATGACTCGGAATATGAAATTGTTGCTTCCAGATGTTCTGGTTCCAGG
ATTCCAAGCAGGTCTGTTCCAGCCTCTGATTGTTGATCGGGCAAGATTGCACAGTACAGT
GTATGAAGTTATTCAAGCACATCCCTGCCAGCAGCAAGACCATCCAGAGTGAACATTG
GCTAAACAGTACATTCAAGTGAAGAAACATTGAAGAAACATTAAAGAAAAACAGTGGAAAAGT
ATATTAACTGGAATCACTGAAAGAACCCAGGACCAACACCTTACTCATTATTCTTACA
TGCAGAATAGAGGCATTATGCAAATTGAACTGCAGGTTTCAGCATATACAAATGTCTT
GTGCAACAGAAAACATGTTGGGAAATTCTCAGTGGAGAGTCGTTCTCATGCTGACGG
GGAGAACGAAAGTGACAGGGGTTCTCATAAGTTGATGAAATATCTCACAAACCTCA
ATTAGTTCTACTCTACACTTCACTATCATCAACACTGAGACTATCCTGTCTCACCTACAAA
TGTGGAAACTTACATTGTTGATTTCTCAGCAGACTTTGTTTATTAAATTGTTATTAGTG
TTAAGAATGCTAAATTATGTTCAATTATTTCCAATTCTATCTGTATTGTTACAA
CAAAGTAATAAGGATGGTTGTCAAAAAACAAACTATGCCTCTTTTTCAATCACC
AGTAGTATTGAGAAGACTTGTGAACACTTAAGGAAATGACTATTAAAGTCTTATTGTTA
TTTTTTCAAGGAAAGATGGATTCAAATAAATTATTCTGTTTGCTTTAAAAAAAAAAAAA

FIGURE 112

MWLKVFTTFLSFATGACSLKVTVPSPHTVHGVRGQALYLPVHYGFHTPASDIQIIWLPERPH
TMPKYLLGSVNKSVPDLEYQHKFTMMPPNASLLINPLQFPDEGNYIVKVNIQGNGLTLSASQ
KIQVTVDDPVTKPVVQIHPPSGAVEYVGNMTLTCHVEGGTRLAYQWLKNGRPVHTSSTYSFS
PQNNTLHIAPVTKEDIGNYSCLVRNPVSEMESDIIMPIIYYGPYGLQVNNSDKGLKVGEVFTV
DLGEAILFDCSADSHPPNTYSWIRRTDNTTYIIKHGPRLEVASEKVAQKTMVDYVCCAYNNIT
GRQDETHFTVIITSVGLEKLAQKGKSLSPLASITGISLFLIISMCLLFLWKKYQPYKVIKQK
LEGRPETEYRKAQTFSGHEDALDDFGIYEFVAFPDVSGVSRIPSRSPASDCVSGQDLHSTV
YEVIQHIPAQQQDHPE

Signal sequence:

amino acids 1-18

Transmembrane domain:

amino acids 341-359

N-glycosylation site.

amino acids 73-77, 92-96, 117-121, 153-157, 189-193, 204-208,
276-280, 308-312

Casein kinase II phosphorylation site.

amino acids 129-133, 198-202, 214-218, 388-392, 426-430, 433-437

Tyrosine kinase phosphorylation site.

amino acids 272-280

N-myristoylation site.

amino acids 15-21, 19-25, 118-124, 163-167, 203-209, 231-237,
239-245

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 7-18

FIGURE 113

GCAAGCGGC~~AA~~**ATGGCGCCCTCCGGGAGTCTGCAGT**TCCTGGCAGTCCTGGTGTGTT
GCTTTGGGTGCTCCCTGGACGCACGGCGCGGAGCAACGTTCCGTCATCACGGACGAGA
ACTGGAGAGAACTGCTGGAAGGAGACTGGATGATAGAATTATGCCCGTGGTGCCTGCT
TGTCAAAATCTCAACCGGAATGGGAAAGTTGCTGAATGGGGAGAAGATCTGAGGTTAA
TATTGCGAAAGTAGATGTCACAGAGCAGCCAGGACTGAGTGGACGGTTATCATAACTGCTC
TTCCTACTATTTATCATTGTAAGATGGTGAATTAGGCGCTATCAGGGTCCAAGGACTAAG
AAGGACTTCATAAACTTTATAAGTGTAAAGAGTGGAAAGAGTATTGAGCCGTTCATCATG
GTTTGGTCCAGGTTCTGTTCTGATGAGTAGTATGTCAGCACTCTTCAGCTATCTATGTGGA
TCAGGACGTGCCATAACTACTTTATTGAAGACCTTGGATTGCCAGTGTGGGGATCATATACT
GTTTTGCTTAGCAACTCTGTTCCGGACTGTTATTAGGACTCTGTATGATATTGTGGC
AGATTGCTTGTCTTCAAAAAGGCGCAGACCACAGCCATACCCATACCCCTCAAAAAAAAT
TATTATCAGAATCTGCACAACCTTGAAAAAAAGTGGAGGAGGAACAAGAGGCGGATGAAGAA
GATGTTTCAGAAGAAGACTGAAAGTAAAGAAGGAACAACAAACAAAGACTTCCACAGAATGC
CATAGACAACGCTCTGGTCCATCATTGCCACAGATAATCCT**AGT**AAATTATAG
TTATCTTAATATTATGATTGATAAAAACAGAAGATTGATCATTGTTGGTTGAAGTG
AACTGTGACTTTTGAAATTGCAGGGTTCACTAGATTGTCAATTAAATTGAAGAGTCTA
CATTCAAGAACATAAAAGCACTAGGTATAAGTTGAAATATGATTAAAGCACAGTATGATG
GTTTAAATAGTCTCTAATTGAAAAACGTGCCAAGCAATAAGATTATGTATATTGT
TTAATAATAACCTATTCAAGTCTGAGTTGAAAATTACATTCCCAAGTATTGCAATT
TGAGGTATTAAAGAAGATTTTAGAGAAAATTCTCATTGATATAATTTCCTCTG
TTTCACTGTGAAAAAAAGAAGATATTCCCATAAATGGAAAGTTGCCATTGTCTCAAG
AAATGTGTATTCAGTGACAATTCTGGTCTTTAGAGGTATATTCCAAATTTCCTGT
ATTTTAGGTTATGCAACTAATAAAACTACCTTACATTAATTACAGTTCTACACA
TGGTAATACAGGATATGCTACTGATTAGGAAGTTTAAGTTCATGGTATTCTTGTGATT
CAACAAAGTTGATTCTCTGTATTCTTACTATGGTTACATTTTTATT
CAAATTGGATGATAATTCTTGAAACATTTTATGTTAGTAAACAGTATTGTT
GTTTCAAACCTGAAGTTACTGAGAGATCCATCAAATTGAACAATCTGTTGAATTAAAATT
TTGCCACTTTTCAGATTTCATCATTCTGCTGAACCTCAACTGAAATTGTTTT
TTCTTTTGATGTGAAGGTGAACATTCTGATTGCTGATGTGAAAAAGCCTGGTA
TTTACATTGAAAATTCAAAGCTTAATATAAAAGTTGCATTCTACTCAGGAAAAG
CATCTCTGTATATGTCTAAATGTATTGTCCTCATATACAGAAAGTCTTAATTGAT
TTTACAGTCTGTAATGCTGATGTTAAAATAATAACATTTTATATTGTTAAAAGACAA
ACTTCATATTATCCTGTGTTCTTCTGACTGGTAATATTGTTGTTGATGTT
GTCAGTAGGATGGAACATTAGTGTATTCTACTCCTAAAGAGCTAGAATACATAGTTT
CACCTTAAAGAAGGGGGAAATCATAAATAACATGAATCAACTGACCATTACGTAGTAGAC
AATTCTGTAATGTCCTTCTTCAAGCCCTCTCCTTGAATTAAAGTACATTACAG
TATCGTAATATAACAGTTCTTAAAGCCCTCTCCTTGAATTAAAGTACATTACAG
AAAGAGTTGGATGTGTAATTGTTGATGCCTAGAAAAATCCTAAGCACAAATAACCT
TTCTAACCACTTCATTAAAGCTGAAAAAA

FIGURE 114

MAPSGSLAVPLAVLVLLLWGPWTHGRRSNVRVITDENWRELLEGDWMIEFYAPWCPACQNL
QPEWESFAEWGEDLEVNIAKVDVTEQPGLSGRFIITALPTIYHCKDGEFRRYQGPRTKKDFI
NFISDKEWKSIEPVSSWFGPGSVLMSMSALFQLSMWIRTCHNYFIEDLGLPVWGSYTVFAL
ATLFSGLLGLCMIFVADCLCPSKRRRPQPYPPSKLLSESAQPLKKVEEEQEADEEDVSE
EEAESKEGTNKDFPQNPAIRQRSLGPSLATDKS

Signal sequence:

amino acids 1-26

Transmembrane domain:

amino acids 182-201

Casein kinase II phosphorylation site.

amino acids 68-72, 119-123, 128-132, 247-251, 257-261

Tyrosine kinase phosphorylation site.

amino acids 107-115

N-myristoylation site.

amino acids 20-26, 192-198

Amidation site.

amino acids 25-29

FIGURE 115

GCGAGTGTCCAGCTCGGGAGACCGTGATAATTGTTAACTAATTCAACAAACGGGACCCCTT
 CTGTGTGCCAGAAACCGCAAGCAGTGTCTAACCCAGTGGGACAGGCGGATTGGAAGAGCGGG
 AAGGTCTGGCCCAGAGCAGTGTGACACTTCCCTCTGTGACC**ATGAAACTCTGGGTGTCTGC**
 ATTGCTGATGGCTGGTTGGTGTCTGAGCTGTGAGGCGAATTCTCACCTCTATTG
 GGCACATGACTGACCTGATTATGCAGAGAAAGAGCTGGTGCAGTCTCTGAAAGAGTACATC
 CTTGTGGAGGAAGCCAAGCTTCAAGATTAAGAGCTGGGCAACAAAATGGAAGCCTGAC
 TAGCAAGTCAGCTGCTGATGCTGAGGGCTACCTGGCTCACCTGTGAATGCCTACAAACTGG
 TGAAGCGGCTAAACACAGACTGGCTCGCTGGAGGACCTGTCTGCAGGACTCAGCTGCA
 GGTTTATGCCAACCTCTGTGAGCGGAGTTCTTCCCCACTGATGAGGACGAGATAGG
 AGCTGCCAAAGCCTGATGAGACTTCAGGACACATACAGGCTGGACCCAGGCACAATTCCA
 GAGGGGAACCTCCAGGAACCAAGTACCAAGGCAATGCTGAGTGTGGATGACTGCTTGGGATG
 GGCGCTCGGCCTACAATGAAGGGACTATTATCATACTGGTGTGGATGGAGCAGGTGCT
 AAAGCAGCTTGATGCCGGGGAGGAGGCCACCAACCAAGTCACAGGTGCTGGACTACCTCA
 GCTATGCTGTCTCCAGTTGGGTGATCTGCACCGTGCCTGGAGCTCACCGCCGCTGCTC
 TCCCTGACCAAGCCACGAACGAGCTGGAGGGAACTGCGGTACTTGAGGAGTTATTGGA
 GGAAGAGAGAGAAAAAACGTTAACAAATCAGACAGAAGCTGAGCTAGCAACCCAGAAGGCA
 TCTATGAGAGGCCGTGGACTACCTGCCTGAGAGGGATGTTACGAGAGCCTGTGCTGGG
 GAGGGTGTCAAACGTACACCCCCTAGACAGAAAGAGGCTTCTGTAGGTACCAACATGGCAA
 CAGGGCCCCACAGCTGCTATTGCCCTTCAAAGAGGAGGAGCAGTGGGACAGCCCGACA
 TCGTCAGGTACTACGATGTCATGTCAGGAAATCGAGAGGATCAAGGAGATCGAAAA
 CCTAAACTGCACGAGCCACCGTCTGTGATCCAAAGACAGGAGTCCTCACTGTGCCAGCTA
 CCGGTTCCAAAGCTCCTGGCTAGAGGAAGATGATGACCTGTTGTGGCCGAGTAAATC
 GTCGGATGCAGCATATCACAGGTTAACAGTAAAGACTGCAGAATTGTTACAGGTTGCAAAT
 TATGGAGTGGGAGGACAGTATGAACCGCACTCGACTTCTCTAGGCGACCTTTGACAGCGG
 CCTCAAAACAGAGGGAAATAGGTTAGCGACGTTCTTAACATGAGTGTAGAAGCTG
 GTGGTGCCACCGTCTCCCTGATCTGGGCTGCAATTGGCTAACAGGGTACAGCTGTG
 TTCTGGTACAACCTCTGCGGAGCGGGAAAGGTGACTACCGAACAGACATGCTGCCG
 TGTGCTGTGGCTGCAAGTGGCTCCAATAAGTGGTCCATGAACGAGGACAGGAGTTCT
 TGAGACCTTGTGGATCAACAGAAGTGT**ACTGACATCCTTTCTGTCTTCCCTGGTC**
 CTTCAGCCCATGTCAACGTGACAGACACCTTGTATGTTCTTGTATGTTCTATCAGGCT
 GATTTTGGAGAAATGAATGTTGCTGGAGCAGAGGAGACCATACTAGGGCGACTCCTGT
 GTGACTGAAGTCCCAGCCCTCCATTCAAGCTGTGCCATCCCTGGCCCCAAGGCTAGGATCA
 AAGTGGCTGCAGCAGAGTTAGCTGTCTAGCGCCTAGCAAGGTGCCTTGTACCTCAGGTGTT
 TTAGGGTGTGAGATGTTCACTGAACCAAAGTTCTGATAACCTGTTACATGTTGTTTAT
 GGCATTCTATCTATTGTGGCTTACCAAAAAATAATGTCCCTACCAAGAAAAAA

FIGURE 116

MKLWVSALLMAWFGVLSCVQAEFFTSIGHMTDLIYAEKELVQSLKEYILVEEAKLSKIKSWA
NKMEALTSKSAADAEGYLAHPVNAYKLVKRLNTDWPALEDLVLQDSAAGFIANLSVQRQFFP
TDEDEIGAAKALMRLQDTYRLDPGTISRGELPGTKYQAMLSVDDCFGMGRSAYNEGDYYHTV
LWMEQVLKQLDAGEEATTKSQVLDYLSYAVFQLGDLHRALELTRRLLSLDPHSHERAGGNLR
YFEQLLEEEREKTLTNQTEAELATPEGIYERPVDPYLPERDVYESLCRGEVKLTPRRQKRLF
CRYHHGNRAPQLLIAPFKEEDEWDSPHIVRYYDMSDEEIERIKEIAKPKLARATVRDPKTG
VLTVASYRVSKSSWLEEDDPVVARNRRMQHITGLTVKTAELLQVANYGVGGQYEPHFDFS
RRPFDSDLKTEGNRLATFLNYMSDVEAGGATVFPDLGAIWPKKGTAVFWYNLLRSGEGDYL
TRHAACPVLVGCKWVSNKWFHERGQEFLRPCGSTEV

Signal sequence:

amino acids 1-17

N-glycosylation site.

amino acids 115-119, 264-268

Glycosaminoglycan attachment site.

amino acids 490-494

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 477-481

Casein kinase II phosphorylation site.

amino acids 43-47, 72-76, 125-129, 151-155, 165-169, 266-270,
346-350, 365-369, 385-389, 457-461, 530-534

Tyrosine kinase phosphorylation site.

amino acids 71-80, 489-496

N-myristoylation site.

amino acids 14-20, 131-137, 171-177, 446-452

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 8-19

Leucine zipper pattern.

amino acids 213-235

FIGURE 117

GCAGTATTGAGTTTACTCCTCTTTAGTGAAGACAGACCATAATCCCAGTGTGAGTGAAATTGATTGT
 TTCATTATTACCGTTTGGCTGGGGTTAGTCCGACACCTCACAGTGAAGAGCAGGCAGAAGGAGTTGTGA
 AGACAGGACAATCTCTGGGGATGCTGGCCTGGAAGCCAGCGGGGCTTGCTCTGTCTTGGCCTCATTGACCC
 CAGGTTCTGGTTAAAACGTAAAGCCTACTACTGGCCTGGTGCCTCAATCCATTGATCCTTGAGGCTGTGCC
 CCTGGGGACCCACCTGGCAGGGCTACCACCATTCGACTGAGCTCCCTGTTGGCTCTGCTGCGGCCAGCGCTTC
 CCCTCATCTAGGGCTGTCTGGGGTGCAGCCTGAGCCTCTGCGGGTTCTGGATCCAGGGGAGGGAGAAG
 ATCCCTGTGTCAGGGCTGTAGGGGAGCGAGGAGGGCCACAGAATCCAGATTGAGAGCTCGGCTAGACCAAAGTG
 ATGAAGACTCAAACCCGGATTGTCCCCTACTACAGGGACCCAAACAAGCCCTACAAGAAGGTGCTCAGGACTC
 GGTACATCCAGACAGAGCTGGGCTCCGTGAGCGGTTGTTGGCTGCCTGACCTCCGAGCTACACTGTCCA
 CTTTGGCCGTGGCTGTGAACCGTACGGTGGCCCATCACTTCCCTCGTTACTCTACTTCACGGCAGCGGGGG
 CCCGGGCTCCAGCAGGGATGCAGGTGGTGTCTATGGGATGAGCGCCCGCCTGGCTCATGTCAGAGACCCCTGC
 GCCACCTCACACACACTTGGGCGACTACGACTGGTCTTCATCATGCAAGGATGACACATATGTGCAGGCC
 CCCGCCTGGCAGCCCCCTGCTGGCACCTCAGCATCAACCAAGACCTGACTTAGGCCGGCAGAGGAGTTATTG
 GCGCAGGCAGCAGGCCGGTACTGTATGGGGTTGGTACCTGTTGTCACGGAGTCTCTGCTGCCTGC
 GCCACATCTGGATGGCTGCCAGGGAGACATTCTCAGTGCCTGACGAGTGGCTGGACGCTGCCTCATTG
 ACTCTGGGCGTGTCTCACAGCACCAAGGGCAGCAGTATGCTCATTGAACGGGAAAGGGGGGGGGGGGGGGGGGG
 ACCCTGAGAAGGAAGGGAGCTGGCTTCTGAGTGCTTCGCGTGCACCCCTGTCCTCGAAGGTACCCCTCATGT
 ACCGGCTCCACAAACGCTTCAGCGCTCTGGAGTTGGAGGGCTTACAGTGAATAGAACAACTGCAGGCTCAGA
 TCGGAACCTGACCGTGTGACCCCCGAAGGGGAGGCAGGGCTGAGCTGGCCGTTGGCTCCCTGCTCCCTCA
 CACACACTCTGCTTGGAGGTGTGGACTACTTCACAGAGCAGCACACCTCTCTGTCAGATGGGG
 CTCCCAAGTGCCCACTACAGGGGCTAGCAGGGCGGACGGTGGTATGCGTTGGAGACTGCCCTGGAGCAGCTCA
 ATCGGCCTATCAGCCCCGCTCGCTTCCAGAAGCAGCAGTGTCAACGGCTATGGCCTTGACCCAGCAC
 GGGCATGGAGTACACCTGGACCTGCTTGGAAATCTACCTATGCTACTGAGGCCACCCGAGTGC
 GGGTCACTGGCTGCTGCCACTGAGCGGGTGGAAATCTACCTATGCTACTGAGGCCACCCGAGTGC
 AGCTGGTGTGCCACTCTGGTGGCTGAAGCTGCTGCAGCCCCGGCTTCTCGAGGGCTTGCAGCCAATGTCC
 TGGAGCCACGAGAACATGCAATTGCTCACCCCTGTTGCTACGGCCACGAGAAGGTGGCGTGGAGCTCCAG
 ACCCATTCTGGGGTGAAGGGCTGCAAGCAGGGGTTAGAGCAGCGTACCCCTGGGACGAGGCTGGCTGGCTCG
 CTGTGCGAGCAGAGGCCCTTCCAGGTGCACTCATGGACGTGGTCTGAAGAAGCACCCTGTTGGACACTCTCT
 TCTTCCCTACCAACCGTGTGGACAAGGCCCTGGGCCGAAGTCTCAACCGCTGTCCATGAATGCCATCTGGCT
 GGCAGGCCCTCTCCAGTCCATTCCAGGGCTCAATCTGCCCTGTCAACCACAGAGATCCCCCAGGGCCCC
 CGGGGGCTGGCCCTGACCCCCCCCCTCCCTGGTGTGACCCCTCCGGGGGCTCTATAGGGGGGAGATTG
 ACCGGCAGGCTCTGGGGAGGGCTGCTTACAACGCTGACTACCTGGCGGCCAGGCCGGCTGGCAGGTGAAC
 TGGCAGGCCAGGAAGAGGAGGAAGGCCCTGGAGGGCTGGAGGTGATGGATGTTCTCCGGTTCTCAGGGCTCC
 ACCTCTTCGGGCCGTAGAGCCAGGGCTGGTGCAGAAGTTCTCCCTGCGAGACTGCAGCCCACGGCTCAGTGAAG
 AACTCTACCACCGCTGCCCTCAGCAACCTGGAGGGGCTAGGGGGCGTGCCTAGCTGCTATGGCTCTTTG
 AGCAGGAGCAGGCCAATAGCACTTCGCCGCTGGGGGCCCTAACCTCATTACCTTCTGTCTGCCCTAGCC
 CCAGGAAGGGCAAGGCAAGATGGTGGACAGATAGAGAATTGTTGCTGTATTTTAAATATGAAAATGTTATTAA
 ACATGTCTCTGCC

FIGURE 118

MRLSSLLALLRPALPLILGLSLGCSLSLLRVSWIQGEGEDPCVEAVGERGGPQNPDZRARLD
QSDEDFKPRIVPYRDPNPKVVKLTRYIQTTELGSRERLLVAVLTSRATLSTLAVAVNRTV
AHHPRLLYFTGQRGARAPAGMQVVSHGDERPAWLMSETRHLHHTFGADYDWFFIMQDDTY
VQAPRLAALAGHLSINQDLYLGRAEEFIGAGEQARYCHGGFGYLLRSLLLRLRPHLDGCRG
DILSARPDEWLGRCLIDSLGVGCVSQHQGQQYRSFELAKNRDPEKEGSSAFLSAFAVHPVSE
GTLMYRLHKRFSALELERAYSEIEQLQAQIRNLTVEPEGEAGLSWPVGLPAPFTPNSRFEV
LGWDYFTEQHTFSCADGPKCPLQGASRADVGALETALEQLNRRYQPRLRFQKQRLLNGYR
RFDPARGMETLDLLLECVTQRGHRRALARRVSLRPLSVEILPMPYVTEATRVLQLVPL
VAEAAAAPAFLEAFAAANVLEPREHALLTLLVYGPREGGRGAPDPFLGVKAAAELERRYPG
TRLAWLAVRAEAPSQVRLMDVVSKKHPVDTLFFLTTVWTRPGPEVLNRCRMNAISGWQAFFP
VHFQEFPALSPQRSPPGAGPDPPSPFGADPSRGAPIGGRFDRQASAEGCFYNADYLAA
RARLAGELAGQEEEAELEGLEVMDVFLRFSGLHLFRAVEPGLVQKFSLRDCSPRLSEELYHR
CRLSNLEGLGGRQLAMALFEQEQANST

Signal sequence:

amino acids 1-15

Transmembrane domain:

amino acids 489-507

N-glycosylation site.

amino acids 121-125, 342-346

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 319-323, 464-468

Casein kinase II phosphorylation site.

amino acids 64-68, 150-154, 322-326, 331-337, 368-372, 385-389,
399-403, 409-413, 473-477, 729-733, 748-752

Tyrosine kinase phosphorylation site.

amino acids 736-743

N-myristoylation site.

amino acids 19-25, 23-29, 136-142, 397-403, 441-447, 544-550,
558-564, 651-657, 657-663, 672-678

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 14-25

Cell attachment sequence.

amino acids 247-250

FIGURE 119

CGGAGTGGTGCGCCAACGTGAGAGGAAACCGTGC CGC GCTGC GCTT CCGT CCCCCAAGCC
GTTCTAGACGCGGGAAAAATGCTTCTGAAAGCAGCTCCTTTGAAGGGTGTGATGCTTGG
AAGCATTCTGTGCTTGATCACTATGCTAGGACACATTAGGATTGGTCATGGAAATAGAA
TGCACCACCATGAGCATCATCACCTACAAGCTCTAACAAAGAAGATATCTGAAAATTCA
GAGGATGAGCGCATGGAGCTCAGTAAGAGCTTCGAGTATACTGTATTATCCTGTAAAACC
CAAAGATGTGAGCTTGGCTGCAGTAAAGGAGACTTGGACCAAACACTGTGACAAAGCAG
AGTTCTCAGTTCTGAAAATGTTAAAGTGTGAGTCATTAATATGGACACAAATGACATG
TGGTTAATGATGAGAAAAGCTTACAAATACGCCCTTGATAAGTATAGAGACCAATACAATG
GTTCTCCTGCACGCCCACTACGTTGCTATCATTGAAAACCTAAAGTATTTTGTAA
AAAAGGATCCATCACAGCCTTCTATCTAGGCCACACTATAAAATCTGGAGACCTTGAATAT
GTGGGTATGGAAGGAGGAATTGTCTTAAGTGTAGAATCAATGAAAAGACTTAACAGCCTCT
CAATATCCCAGAAAAGTGTCTGAACAGGGAGGGATGATTGGAGAGATATCTGAAGATAAAC
AGCTAGCAGTTGCCTGAAATATGCTGGAGTATTGAGAAAATGCAGAAGATGCTGATGGA
AAAGATGTATTAATACAAATCTGTTGGCTTCTATTAAAGAGGCAATGACTTACACCC
CAACCAGGTAGTAGAAGGCTGTTAGATATGGCTGTTACTTTAATGGACTGACTCCAA
ATCAGATGCATGTGATGTATGGGTATACGCCCTAGGGCATTGGCATATTTCAAT
GATGCATTGGTTTCTTACCTCAAATGGCTGACAATGACTTGAGAAGTGGTAGAAAAGCG
TGAATATGATCTTGTATAGGACGTGTGTCATTATTGAGTAGTAACATATCCAA
TACAGCTGTATGTTCTTTCTTAATTGGCTGACTGGTATAACCACACATTAAAG
TCAGTAGTACATTTAAATGAGGGGGTTTTCTTAAACACATGAACATTGAAATG
TGTGGAAAGAAGTGTGTTAAGAATAATAATTGCAAAACTATTAAATAATTATAT
GTGATAAAATTCTAAATTGAACATTAGAAATCTGTGGGGCACATATTTGCTGATTGGTT
AAAAAATTAAACAGGTCTTAGCCTCTAAGATATGCAAATGATATCTCTAGTTGTGAATT
TGTGATTAAAGTAAAACCTTGTGTTCCCTTACTCTAAACTGATTATGTTCT
AAGCCTCCCCAAGTCCAATGGATTGCCTCTCAAAATGTACAACTAAGCAACTAAAGAAA
ATTAAAGTGAAGTTGAAAAAT

FIGURE 120

MLSESSSFLKGVMLGSIFCALITMLGHIRIGHGNRMHHHEHHHLQAPNKEDILKISEDERME
LSKSFRVYCIILVKPKDVSLWAAVKETWTKHCDKAFFSSENVKFESINMDTNDMWLMMRK
AYKYAFDKYRDQYNWFFLARPTTFAIIENLKYFLLKKDPSQPFYLGHТИKSGDLEYVGMEGG
IVLSVESMKRLNSLLNIPEKCPEQGGMIWKISEDKQLAVCLKYAGVFAENAEDADGKDVFNT
KSVGLSIKEAMTYHPNQVVEGCCSDMAVTFNGLTPNQMHVMMYGVYRLRAFGHIFNDALVFL
PPNGSDND

Signal sequence:

amino acids 1-33

N-glycosylation site.

amino acids 121-125, 342-346

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 319-323, 464-468

Casein kinase II phosphorylation site.

amino acids 64-132, 150-154, 322-326, 331-335, 368-372, 385-389,
399-403, 409-413, 473-477, 729-733, 748-752

Tyrosine kinase phosphorylation site.

amino acids 736-743

N-myristoylation site.

amino acids 19-25, 23-29, 136-142, 397-403, 441-447, 544-550,
558-564, 651-657, 657-663, 672-672

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 14-25

Cell attachment sequence.

amino acids 247-250

FIGURE 121

CCACACGCGTCCGATCTTACCAACAAAACACTCCTGAGGAGAAAGAAAGAGAGGGAGGGAGAG
AAAAAGAGAGAGAGAGAAACAAAAACCAAAGAGAGAGAAAAATGAATTCAATCTAAATCAT
CTGAAACACAATGCACAGAGAGAGGATGCTTCTTCCAAATGTTCTTATGGACTGTTGCT
GGGATCCCCATCCTATTCTCAGTGCCTGTTCATCACCAGATGTGTTGACATTCGCAT
CTTTCAAACCTGTGATGAGAAAAAGTTCAGCTACCTGAGAATTACAGAGCTCTCCTGCT
ACAATTATGGATCAGGTTCAAGAATTGTTGTCATTGAACGGAAATTTCACAGAGCTCTCCTGCT
AGCTGCTACTTCTTTCTACTGACACCATTCCCTGGCGTTAAGTTAAAGAACTGCTCAGC
CATGGGGCTCACCTGGTGGTTATCAACTCACAGGAGGAGCAGGAATTCCCTTACAAGA
AACCTAAAATGAGAGAGTTTTTATTGGACTGTCAGACCAGGTTGTCAGGGTCAGTGGCAA
TGGGTGGACGGCACACCTTGACAAAGTCTGAGCTTCTGGATGTAGGGAGGCCAACAA
CATAGCTACCCCTGGAGGACTGTGCCACCAGAGACTCTCAAACCCAAGGCAAATTGGA
ATGATGTAACCTGTTCTCAATTATTCGGATTGTGAAATGGTAGGAATAATCCTTG
AACAAAGGAAAATCTCTTAAGAACAGAACAGGACAACTCAAATGTGAAAGGAAGGAGCA
AGAACATGGCCACACCCACCGCCCCACACGAGAAATTGTCGCTGAACTTCAAAGGACTTC
ATAAGTATTGTTACTCTGATAAAATAAGTAGTTAAATGTTAAAAAAAAAAAAAAA
AAA
AAAAA

FIGURE 122

MNSSKSSETQCTERGCFSSQMFLWTVAGIPIFLSACFITRCVVTFRIFQTCDEKKFQLPEN
FTELSCNYGSGSVKNCCPLNWEYFQSSCYFFSTDТИSWALSLKNCSAMGAHLVVINSQEEQ
EFLSYKKPKMREFFIGLSDQVVEGQWQWVDTPLTKSLSFWDVGEPPNNIATLEDCATMRDSS
NPRQNWNDVTCFLNYFRICEMVGINPLNKGKSL

Signal sequence:

amino acids 1-42

N-glycosylation site.

amino acids 2-6, 62-66, 107-111

Casein kinase II phosphorylation site.

amino acids 51-55, 120-124, 163-167, 175-179, 181-185

N-myristoylation site.

amino acids 15-21, 74-80, 155-161

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 27-38